MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LINCOLN LABORATORY

ADAPTIVE DETECTION IN NON-STATIONARY INTERFERENCE, PART III

E.J. KELLY
Group 96

TECHNICAL REPORT 761

24 AUGUST 1987

LEXINGTON   MASSACHUSETTS

ABSTRACT

The analysis of Parts 1 and 11 of the report with this title has been extended in two directions. In the first case, the performance of an adaptive system with respect to signals arriving from directions other than the steering direction is evaluated. It is shown that these signals are rejected more strongly than would be suggested by the sidelobe levels of the adaptive patterns themselves. In the other case, the detection problem is generalized to include the detection of signals known only to lie in a subspace of the space of steering vectors. Again, performance is derived and the penalty associated with the greater uncertainty of the signal model is shown to be small. The analysis of Part I is essentially repeated here, both to keep this report self-contained and to present an alternative version of the basic derivations.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

1. INTRODUCTION

This study consists of two separate generalizations of an earlier analysis[1] of an adaptive detection algorithm. The problem studied in Reference 1 is very general and easily described. A physical system, such as an antenna array, provides a discrete set of vector-valued data samples. The observation space is N dimensional, and a typical sample vector is called z. This system is plagued with interference, which is modeled as Gaussian noise with zero mean. The N × N covariance matrix of this Gaussian interference is completely unknown, and must be estimated from the data in some way. The problem is to detect the presence of a signal in one data sample, called the "primary" sample, under the assumption that K additional, independent, "secondary" samples are available which are signal-free. The same unknown covariance matrix is assumed to be shared by all these samples, and the signal which is sought is a vector of known direction (the "steering direction") in the observation space, but with an unknown complex amplitude.

The chief difference between the problem thus formulated and the more familiar problem of interference rejection lies in the fact that we are seeking a decision rule for target presence, rather than a filter which will reduce the interference before the data are passed on for further processing and eventual detection. A decision algorithm was derived in Reference 1 by application of the maximum-likelihood principle, and it turns out that the conventional adapted weighting of the sample vector, usually associated with the nulling of interference, appears here as a preliminary step in the detection process. The complete algorithm provides a constant false alarm rate (CFAR) detector which is completely independent of the actual covariance matrix of the interference. The performance of this algorithm was also obtained in Reference 1, in terms of the probabilities of detection and false alarm, and numerical results were presented.

In order to carry out effective detection with a single primary data vector, it is generally necessary that adequate coherent processing be applied to each of the elements of these vectors before detection is attempted. For very large arrays, this of course implies a significant replication of the coherent portions of the processor. One example, which is currently under study, is a radar in which a conventional phased array antenna is used for transmission, and in which coherent (Doppler) processing is applied to each array element, or to a suitable set of subarray outputs, before application of the adaptive detection algorithm to the received signals.

In the first of the generalizations included here, the performance of this same algorithm is obtained for the "mismatched" case of a signal which arrives from a direction different from that for which the system was steered. In general it is desirable that such a signal be rejected, along with the interference, although its presence was not included in the formulation of the hypotheses for target detection. We have in mind a system which must search out a surveillance volume of some kind, and we want the system response to a given signal to fall off rapidly as we steer away from its direction.

It is known that fully adapted arrays often have undesirable patterns in the sidelobes, especially if all their degrees of freedom are absorbed in the task of interference rejection. This sidelobe response has been studied by Boroson[2] in terms of the signal to noise ratio (SNR) developed on such a target by a conventional fully adapted array. Boroson's work generalizes the analysis of Reed, Mallett and Brennan[3] on the SNR associated with a signal arriving from the hypothesized direction in a fully adapted array, and it appears from his analysis that sidelobe response can be controlled only by the use of a very large number of secondary samples for covariance estimation.

The behavior of an adaptive detection system which uses the algorithm discussed above is quite different, and it turns out that the CFAR feature of this algorithm also causes the system response (as measured by its detection probability) to be much less sensitive to mismatched signals. In a sense, the component of a signal which is orthogonal to the steering vector (in the observation space) helps to raise the decision threshold, thereby reducing the probability of its detection.

The analysis of performance for this seemingly minor variation of the problem actually entails considerable complication. It is presented in full in this study, which overlaps and extends the analysis of Reference 1. The present discussion is complete in itself, including some details which were not given in the earlier report. We would like to evaluate the performance of this decision rule in a more general mismatched situation, especially one in which signal components of some kind are present in the secondary data vectors as well. This remains a goal, and as yet only minor steps in that directon have been successfully taken.

The second generalization included here is rather easier, although it is perhaps of less practical interest. The signal hypothesis itself is generalized, so that we are testing for the presence of a signal which is hypothesized only to lie in some definite subspace of the observation space. The interference is modeled the same way as before, and the same input data are assumed to be available for decision. The examples which have so far come to mind are rather artificial, but in view of the generality of

the model, and the fascination (for some) of the mathematics, this case has been included here.

The decision rule turns out to be a kind of noncoherent integrator, which tests for signals in each of the basic directions of the subspace in turn, and combines the absolute squares of the results of these tests. If the subspace is, in fact, the entire space, then one is testing for the presence of a deterministic signal of arbitrary direction and amplitude in the observation space, and also minimizing the response to interference of unknown character. Numerical results are given for the performance of this detector, as well as for the case of the mismatched signal discussed above.

Section 2 provides some background for the mismatch problem, including a statement of Boroson's results. The performance analysis for the mismatched case occupies Sections 3, 4, and 5, and these are mathematically quite detailed. Sections 3 and 4 cover the same ground as Reference 1, but with a slightly different emphasis which may make this derivation somewhat more direct. The analysis of the probability density function (PDF) of the so-called loss factor is given in Section 5, and this material is new. Section 5 also includes a derivation of the detection probability which uses the methods of contour integration, and this material supplements that of Reference 1 where the basic result was quoted without proof.

A discussion of the results for a mismatched signal, with curves and few formulas, appears in Section 6.

Section 7 is devoted to the generalization of the signal model, building on some material from Section 5, and the results in this area are discussed in terms of performance curves in Section 8. The numerical analysis necessary to obtain real answers is treated rather briefly in the Appendix.

2. PROBLEM FORMULATION AND BACKGROUND

Consider an adaptive antenna system with N output channels. The channel inputs are connected to an arbitrary collection of array elements, subarrays, or beamformer networks. In the cases of subarrays and beamformers, the weightings used to produce these inputs remain constant during the processing period considered in this discussion. A single time sample of the antenna output is described by a vector z in a complex N dimensional space:

$$z = [z_1, \ldots, z_N]^T ,$$

where the superscript T stands for matrix transpose.

The possible antenna outputs due to individual sources, in the absence of any noise or interference, comprise a family of vectors which will be written in the form

$$b s_\alpha ,$$

where b is a complex scalar amplitude and α is a directional, or source locational, parameter. These signal vectors are normalized to unity, as follows:

$$(s_\alpha^\dagger s_\alpha) = 1 ,$$

where the symbol $\dagger$ denotes Hermitian conjugate. The dimensionality of α depends upon the system being modeled. For example, for a simple linear array α would be one dimensional, while for a two-dimensional array focusing on near-field as well as far-field sources, it would be a three-dimensional parameter.

The total interference, including the contribution of system noise, is assumed to be Gaussian with zero mean and with covariance matrix M, which is generally unknown. This covariance will be assumed to be constant over the adaptation and processing period of interest here. The matrix M will ultimately be estimated from the data itself, by making use of K additional samples of the antenna output vector, which we refer to as "secondary samples." For the moment, however, we consider the matrix M to be known, and we refer to properties of the system evaluated with this known M as "asymptotic," since they correspond formally to the limit K → ∞.

With known M, the optimum processor for the detection of the signal bs_α for a given parameter value α, is, of course, the matched filter. This filter forms the scalar quantity

$$\xi \equiv (w_\alpha^\dagger z) \; ,$$

where the weight vector is given by

$$w_\alpha = k\, M^{-1} s_\alpha \; .$$

Here, k is an arbitrary constant and detection is based on the magnitude of ξ. The choice of k and the decision threshold is based on the required probability of false alarm (PFA). The signal to noise ratio (SNR) for this detector, when the desired signal is actually present with amplitude b, is given by the well known expression

$$\text{SNR}_{\alpha,\alpha} = |b|^2 \, (s_\alpha^\dagger M^{-1} s_\alpha) \; . \tag{2-1}$$

We use the term "noise" now to refer to the total of interference and system noise.

If, however, the data vector z contains a signal with some other directional parameter value (such as β) when the antenna is "steered" as above for the value α, then the resulting SNR will be

$$\text{SNR}_{\alpha,\beta} = |b|^2 \, \frac{|(s_\alpha^\dagger M^{-1} s_\beta)|^2}{(s_\alpha^\dagger M^{-1} s_\alpha)} \; , \tag{2-2}$$

assuming the signal amplitude is again b. Since the inverse of M is positive definite,

$$(A^\dagger M^{-1} B)$$

may be interpreted as an inner product of A and B. We can therefore write

$$\frac{(s_\alpha^\dagger M^{-1} s_\beta)}{(s_\beta^\dagger M^{-1} s_\beta)^{1/2} (s_\alpha^\dagger M^{-1} s_\alpha)^{1/2}} \equiv \cos\Theta \, e^{i\varphi} \; , \tag{2-3}$$

where Θ is a real angle in the range $0 \le \Theta \le \pi/2$.

The SNR corresponding to our general "sidelobe signal" can now be expressed in the form

$$\mathrm{SNR}_{\alpha,\beta} = \mathrm{SNR}_{\beta,\beta} \cos^2 \Theta . \tag{2-4}$$

The factor $\mathrm{SNR}_{\beta,\beta}$, which represents the optimum SNR which would be attained if the antenna system was steered in the β direction, is reduced by the second factor $\cos^2 \Theta$ which therefore describes the asymptotic sidelobe response of the antenna when steered in the α direction. The actual gain of the antenna in the β direction when steered for α is proportional to the quantity

$$\frac{|(w_\alpha^\dagger s_\beta)|^2}{(w_\alpha^\dagger w_\alpha)} = \frac{|(s_\alpha^\dagger M^{-1} s_\beta)|^2}{(s_\alpha^\dagger M^{-2} s_\alpha)} ,$$

and hence null values of $\mathrm{SNR}_{\alpha,\beta}$ correspond to actual nulls of this asymptotic antenna pattern.

Another case of interest here corresponds to the use of a steering vector different from s_α in order to detect a signal in the α direction. For example, we might use the weight vector

$$w = k M^{-1} q ,$$

where q is a vector obtained by multiplying s_α by a real diagonal matrix whose elements correspond to some chosen antenna taper. When the interference consists only of system noise (the quiescent case), the antenna pattern will then revert to the desired pattern, which corresponds to this taper. This technique is used to lower antenna sidelobes in interference-free regions, and is generally satisfactory if the system has sufficient degrees of freedom with respect to the interference actually present. The use of a tapered steering vector results in a SNR for an actual signal in the α direction which is given by

$$\mathrm{SNR}_{\alpha,\alpha} \cos^2 \Theta ,$$

where in this case

7

$$\cos^2\Theta = \frac{|(s_\alpha^\dagger M^{-1} q)|^2}{(s_\alpha^\dagger M^{-1} s_\alpha)(q^\dagger M^{-1} q)} \ .$$

The second factor now accounts for the performance loss which results from the deliberate departure from the matched filter weights. A sidelobe canceling system provides another example of this kind, since the signal components in the auxiliary channels are usually ignored, resulting in a slightly mismatched steering vector.

Now suppose that the noise covariance matrix M is unknown and must be estimated from a set of K "secondary data samples" $z(k)$. We assume that these samples are independent and statistically identical, sharing the unknown covariance matrix, and that they are all zero mean Gaussian vectors. In radar practice, these samples usually correspond to range bins other than the "primary" one in which detection is to be carried out. We assume now that the sample covariance matrix is used as an estimator of M, writing

$$\widehat{M} = \frac{1}{K} S \ ,$$

where

$$S \equiv \sum_{k=1}^{K} z(k) z(k)^\dagger \ .$$

The matrix S is subject to the Wishart probability density function.

We now choose

$$w_\alpha = S^{-1} s_\alpha$$

as an adapted weight vector, using the conventional steering vector s_α. For given values of the secondary data vectors, we may consider the SNR which results when this weight vector is used with the primary sample z. If the latter actually contains a signal corresponding to the β direction, and if the actual interference covariance matrix is M, this SNR will be

$$\text{SNR} = |b|^2 \frac{|(w_\alpha^\dagger s_\beta)|^2}{(w_\alpha^\dagger M w_\alpha)} = |b|^2 \frac{|(s_\alpha^\dagger S^{-1} s_\beta)|^2}{(s_\alpha^\dagger S^{-1} M S^{-1} s_\alpha)} \ .$$

In analogy to Equation (2-4) we write

$$\text{SNR} = \text{SNR}_{\beta,\beta} \, \rho_{\alpha,\beta} \ , \tag{2-5}$$

and we interpret the quantity

$$\rho_{\alpha,\beta} = \frac{|(s_\alpha^\dagger S^{-1} s_\beta)|^2}{(s_\alpha^\dagger S^{-1} M S^{-1} s_\alpha)(s_\beta^\dagger M^{-1} s_\beta)} \tag{2-6}$$

as a SNR "loss" relative to $\text{SNR}_{\beta,\beta}$, the maximum asymptotic SNR in the β direction. If we were to put

$$M^{-1} s_\beta \equiv v \ ,$$

we would have

$$\rho_{\alpha,\beta} = \frac{|(w_\alpha^\dagger M v)|^2}{(w_\alpha^\dagger M w_\alpha)(v^\dagger M v)} \ ,$$

which shows that this loss factor lies in the range [0,1], again by the use of the Schwarz inequality. It is not difficult to show that the value unity can actually be attained, for a realizable value of the sample covariance matrix. Therefore the SNR can actually be much larger, when M is being estimated, than the limiting value

$$\rho_{\alpha,\beta} \xrightarrow[K \to \infty]{} \cos^2 \Theta \ ,$$

for signals arriving in the low sidelobe region of the asymptotic antenna pattern.

Considered as a function of S, $\rho_{\alpha,\beta}$ is a random variable – a direct generalization of the loss factor obtained by Reed, Mallett, and Brennan[3] – to which it reduces when $\beta = \alpha$. The probability density function (PDF) of $\rho_{\alpha,\beta}$ can be explicitly obtained by a rather tedious computation, and the result can be found in Reference 2. In our notation, Boroson's formula for this PDF can be expressed in the compact form

9

$$f(\rho;\Theta) = \sum_{m=0}^{L} \binom{L}{m} \left(\cos^2\Theta\right)^{L-m} \left(\sin^2\Theta\right)^m f_\beta(\rho;L+1-m,N-1+m) , \qquad (2\text{-}7)$$

where Θ is defined as before, using the actual covariance matrix and signal vectors;

$$L \equiv K+1-N \qquad (2\text{-}8)$$

and f_β is the Beta PDF:

$$f_\beta(x;n,m) = \frac{(n+m-1)!}{(n-1)!\,(m-1)!} \, x^{n-1}(1-x)^{m-1} . \qquad (2\text{-}9)$$

Boroson's parameter d corresponds to $\sec^2\Theta$, in our notation, and his random variable ρ_2 is equal to $d\rho$. Recall that N is the dimension of all data vectors and that K is the number of secondary samples.

The case $\Theta = 0$ corresponds to a perfect match (i.e., $s_\beta = s_\alpha$) and in this case the loss factor PDF reduces to

$$f(\rho;0) = f_\beta(\rho;L+1,N-1) ,$$

which is exactly the result of Reed, Mallett, and Brennan. The other extreme, when $\Theta = \pi/2$, represents the case of a signal which is in a null of the asymptotic pattern, and in this situation we find the simple formula

$$f(\rho;\pi/2) = f_\beta(\rho;1,N+L-1) = K(1-\rho)^{K-1} .$$

The mean value of ρ in this case is

$$\bar{\rho}_{\alpha,\beta} = \frac{1}{K+1} ,$$

which illustrates the slow decline (with increasing K) of the sidelobe SNR, even for a signal in a null of the asymptotic pattern. This behavior should be typical of signals in the low-sidelobe region of the asymptotic antenna pattern as a whole. In general, the PDF of the loss factor is intermediate between these extremes, and its mean value is given by the general formula

$$\bar{\rho}_{\alpha,\beta} = \frac{1}{K+1}\left[1 + (K+1-N)\cos^2\Theta\right],$$

which converges to the asymptotic value with increasing K. In addition, for large K the standard deviation of the loss factor varies inversely with K. It should be clear that the loss factor PDF applies equally well to the case of the mismatched steering vector q, since it is a function only of the angle Θ and the integer parameters N and K.

3. THE LIKELIHOOD RATIO TEST

In Section 2 we discussed the SNR developed when a single sample of an antenna array output vector is processed, using a weight vector derived from an additional set of signal-free samples. In many applications, target detection will be based on the weighted outputs corresponding to a group of samples, such as the radar returns associated with a given range bin for an entire coherent train of pulses. The weight vector may be updated at intervals or remain unchanged throughout the coherent processing interval. If, however, the decision on target presence is based on a single output sample (but still using a weight vector derived from additional, secondary data), then a more complete analysis is possible. It is an interesting result of this analysis that such a system is much less vulnerable to signals in the sidelobe region than the SNR analysis alone suggests.

In order to base target detection on a single adapted antenna sample, adequate coherent processing will need to have been carried out on the individual antenna channel outputs themselves. An example would be a radar in which Doppler processing is performed in every channel, followed by adaptive beamforming for each Doppler frequency.

Instead of separating the problem into a covariance matrix estimation procedure, using the K secondary samples, followed by a matched filtering operation (i.e., adaptive beamforming) on the single primary sample, it can now be formulated as a decision problem using all K + 1 data samples. These samples are assumed to be independent Gaussian vectors, sharing a common, unknown covariance matrix M. The K secondaries are taken to be zero mean (signal-free) vectors, and the primary sample may or may not contain a signal, characterized by a steering vector such as bs_α. A likelihood ratio (LR) decision rule appropriate to this problem was derived in Reference 1, and its performance in terms of probability of detection (PD) and PFA was evaluated for the case in which the primary vector actually contains the postulated signal with some complex amplitude b. In the present study, this analysis is generalized so that the steering vector and the actual signal component of the primary vector are arbitrary and may be different. The case of a deliberately mismatched steering vector is thus automatically included here, as well as the sidelobe response of the system when a conventional steering vector is used.

In the notation of Section 2, the primary sample is z and the secondary samples are z(k), where $1 \leq k \leq K$, and all these data samples are N dimensional complex vectors. The postulated signal vector will be called q. It is normalized to unity: $(q^\dagger q) = 1$,

13

and it will be tacitly assumed that q is some permissible true signal vector such as s_α, or a tapered version of such a vector. The LR decision rule, or criterion for target presence, is then

$$l \geq l_0 ,$$

where l_0 is a threshold parameter and l is given by

$$l = \frac{1 + (z^\dagger S^{-1} z)}{1 + (z^\dagger S^{-1} z) - \frac{|(q^\dagger S^{-1} z)|^2}{(q^\dagger S^{-1} q)}} . \tag{3-1}$$

As shown in Reference 1, l is the $(K+1)^{st}$ root of the maximized likelihood ratio for this problem.

In this expression, S is again the sum

$$S \equiv \sum_{k=1}^{K} z(k) z(k)^\dagger . \tag{3-2}$$

The secondary data enter the decision rule only through S, which is K times the sample covariance matrix of these secondaries. This structure is a consequence of the LR test formulation, and is not an *a priori* choice, as it was in the discussion of Section 2. We note also that the normalization of the steering vector q has no effect on the form of the LR test.

It was shown in Reference 1 that the LR test is a true CFAR (constant false alarm rate) decision rule, whose performance in the absence of any signal components is totally independent of the actual covariance matrix M which describes the interference. We now seek to evaluate the PD of this test when the primary vector contains an arbitrary signal component, which we take to be

$$E z = b p . \tag{3-3}$$

The vector p is also normalized to unity: $(p^\dagger p) = 1$, and b is a complex scalar amplitude. It is assumed that p is some permissible signal vector, such as s_β, but the analysis depends only on p, q and the actual covariance matrix M. We also introduce the notation

14

$$A_q^2 \equiv (q^\dagger M^{-1} q)$$

$$A_p^2 \equiv (p^\dagger M^{-1} p) \tag{3-4}$$

and

$$\cos\Theta \, e^{i\varphi} \equiv \frac{(q^\dagger M^{-1} p)}{A_q A_p} \; . \tag{3-5}$$

The relationship of these parameters to the asymptotic SNR values and sidelobe levels was discussed in Section 2.

The performance analysis is facilitated by a sequence of changes of coordinates, first whitening and then rotating the vectors. Part of this analysis is identical to that given in Reference 1, but it will be presented here in full, both for completeness and to illustrate some variations in the method. First, a positive definite square root of the actual covariance matrix is chosen, and new, whitened coordinates are introduced as follows:

$$\bar{z} \equiv M^{-1/2} z$$

$$\bar{z}(k) \equiv M^{-1/2} z(k) \; , \qquad 1 \leq k \leq K \; .$$

Then

$$\bar{S} \equiv M^{-1/2} S M^{-1/2}$$

is K times the sample covariance matrix of the whitened secondaries:

$$\bar{S} = \sum_{k=1}^{K} \bar{z}(k) \, \bar{z}(k)^\dagger \; .$$

The covariance matrix of the whitened vectors is now the $N \times N$ identity, and the expected value of the whitened primary vector is

$$E\bar{z} = bM^{-1/2} p \; .$$

We note that

15

$$|E\bar{z}|^2 = |b|^2 A_p^2 \ ,$$

and also that

$$A_q^2 = (\bar{q}^{\dagger}\bar{q}) \ ,$$

where

$$\bar{q} \equiv M^{-1/2}q$$

is the whitened steering vector.

By direct substitution it is easily verified that the likelihood ratio is unchanged in form:

$$\ell = \frac{1 + (\bar{z}^{\dagger}\bar{S}^{-1}\bar{z})}{1 + (\bar{z}^{\dagger}\bar{S}^{-1}\bar{z}) - \dfrac{|(\bar{q}^{\dagger}\bar{S}^{-1}\bar{z})|^2}{(\bar{q}^{\dagger}\bar{S}^{-1}\bar{q})}} \ . \tag{3-6}$$

Since this expression is independent of the normalization of the steering vector, we can replace $\bar{q}$ by the unit vector $\bar{e}$, which is defined by

$$\bar{e} \equiv \frac{\bar{q}}{A_q} \ .$$

The next step is another change of coordinates, by means of a unitary transformation U, chosen so that the unit vector $\bar{e}$ is transformed into the basis vector e:

$$e = U\bar{e} = [1,0,...,0]^T \ .$$

To simplify the writing, the new data vectors are represented by the original symbols. This should not cause serious confusion, since we have no need to return to the data vectors in the original coordinates, and we require only a statistical characterization of the quantities entering the LR decision rule in order to evaluate its performance. Thus we write now

$$\mathbf{z} \equiv U\bar{z}$$

$$z(k) \equiv U\bar{z}(k)$$

and

$$S \equiv U\bar{S}U^\dagger = \sum_{k=1}^{K} z(k)\, z(k)^\dagger \ .$$

Then the LR test itself is

$$\ell = \frac{1 + (z^\dagger S^{-1} z)}{1 + (z^\dagger S^{-1} z) - \dfrac{|(e^\dagger S^{-1} z)|^2}{(e^\dagger S^{-1} e)}} \ . \tag{3-7}$$

The new data vectors are independent Gaussian variables, all sharing the identity matrix as a covariance, and the secondaries, of course, have zero mean. Only the expected value of the primary vector is affected by the unitary transformation, and we now have

$$E\,z = bUM^{-1/2}p \ .$$

The norm of this vector is unchanged by the unitary transformation, and we can therefore write

$$E\,z = bA_p f \ ,$$

where f is the unit vector:

$$f \equiv \frac{1}{A_p} UM^{-1/2}p \ .$$

We now decompose f into a component parallel to e and an orthogonal remainder. We compute the inner product of e and f, working back through the transformations and using the various definitions introduced along the way:

17

$$(e^\dagger f) = \frac{1}{A_p}(e^\dagger U M^{-1/2} p)$$

$$= \frac{1}{A_p}(\bar{e}^\dagger M^{-1/2} p) = \frac{1}{A_p A_q}(\bar{q}^\dagger M^{-1/2} p)$$

$$= \frac{1}{A_p A_q}(q^\dagger M^{-1} p) = \cos\Theta e^{i\varphi} .$$

The norm of the component of f which is orthogonal to e is, of course, $\sin\Theta$, since

$$[f - (e^\dagger f) e]^\dagger [f - (e^\dagger f) e] = 1 - \cos^2\Theta = \sin^2\Theta .$$

Finally, we can express the desired decomposition in the form

$$f = \cos\Theta e^{i\varphi} e + \sin\Theta g , \tag{3-8}$$

where g is a unit vector orthogonal to e.

Next, we decompose all of the data vectors into two components: an A component consisting of the first element, and an (N−1) dimensional B component, consisting of the rest of the vector. Thus

$$z \equiv \begin{bmatrix} z_A \\ z_B \end{bmatrix}$$

and

$$z(k) \equiv \begin{bmatrix} z_A(k) \\ z_B(k) \end{bmatrix} , \qquad 1 \le k \le K .$$

The steering vector takes the simple form

$$e = \begin{bmatrix} 1 \\ 0 \end{bmatrix} ,$$

and the orthogonal component of f can be written

18

$$g = \begin{bmatrix} 0 \\ h \end{bmatrix} ,$$

where h is a unit vector in the $(N-1)$ dimensional subspace. The expected value of the primary data vector can therefore be expressed as

$$E\,z = b\,A_p \begin{bmatrix} \cos\Theta e^{i\varphi} \\ \sin\Theta h \end{bmatrix} . \qquad (3\text{-}9)$$

The matrix S and its inverse, which we denote by P, are also decomposed:

$$S = \begin{bmatrix} S_{AA} & S_{AB} \\ S_{BA} & S_{BB} \end{bmatrix} \qquad (3\text{-}10)$$

and

$$P \equiv S^{-1} = \begin{bmatrix} P_{AA} & P_{AB} \\ P_{BA} & P_{BB} \end{bmatrix} . \qquad (3\text{-}11)$$

Since these matrices are Hermitian, we note that the scalars S_{AA} and P_{AA} are real. Also, the matrices S_{BB} and P_{BB} are Hermitian, and

$$S_{BA} = S_{AB}^{\dagger}$$

$$P_{BA} = P_{AB}^{\dagger} .$$

Using these definitions and properties, we can make the evaluation

$$(z^{\dagger} S^{-1} z) = (z^{\dagger} P z)$$

$$= z_A^{\dagger} P_{AA} z_A + z_A^{\dagger} P_{AB} z_B + z_B^{\dagger} P_{BA} z_A + z_B^{\dagger} P_{BB} z_B$$

$$= (z_A + P_{AA}^{-1} P_{AB} z_B)^{*} P_{AA} (z_A + P_{AA}^{-1} P_{AB} z_B)$$

$$+ z_B^{\dagger} (P_{BB} - P_{BA} P_{AA}^{-1} P_{AB}) z_B .$$

The Schur relations for partitioned matrices, applied to S^{-1} and P^{-1}, tell us that

$$P_{AB} = -P_{AA}S_{AB}S_{BB}^{-1} \qquad (3\text{-}12)$$

and also that

$$P_{BB} - P_{BA}P_{AA}^{-1}P_{AB} = S_{BB}^{-1} . \qquad (3\text{-}13)$$

It follows that

$$(z^{\dagger}S^{-1}z) = (z_B^{\dagger}S_{BB}^{-1}z_B) + P_{AA}|z_A - S_{AB}S_{BB}^{-1}z_B|^2 . \qquad (3\text{-}14)$$

The Schur relations also give us an expression for P_{AA}^{-1}, which we denote by T:

$$T \equiv P_{AA}^{-1} = S_{AA} - S_{AB}S_{BB}^{-1}S_{BA} . \qquad (3\text{-}15)$$

We also introduce the notation

$$y \equiv z_A - S_{AB}S_{BB}^{-1}z_B , \qquad (3\text{-}16)$$

which allows us to write

$$(z^{\dagger}S^{-1}z) = (z_B^{\dagger}S_{BB}^{-1}z_B) + \frac{|y|^2}{T} . \qquad (3\text{-}17)$$

The other inner products which enter in the expression for the LR decision rule are easily evaluated:

$$(e^{\dagger}S^{-1}e) = (e^{\dagger}Pe) = P_{AA} = \frac{1}{T} \qquad (3\text{-}18)$$

and

$$(e^{\dagger}S^{-1}z) = (e^{\dagger}Pz) = P_{AA}z_A + P_{AB}z_B$$

$$= P_{AA}(z_A + P_{AA}^{-1}P_{AB}z_B) = \frac{y}{T} .$$

Thus

$$\frac{|(e^{\dagger} S^{-1} z)|^2}{(e^{\dagger} S^{-1} e)} = \frac{|y|^2}{T} \; ,$$

and we have shown that

$$1 + (z^{\dagger} S^{-1} z) - \frac{|(e^{\dagger} S^{-1} z)|^2}{(e^{\dagger} S^{-1} e)} = 1 + (z_B^{\dagger} S_{BB}^{-1} z_B) \; . \tag{3-19}$$

It should be pointed out that the expression for T given by the left side of Equation (3-18) shows that T is K times the maximum likelihood estimator of the spectral density of the interference in the direction specified by the steering vector e.

The results obtained above allow us to write the LR test in the interesting form

$$\ell = \frac{1 + (z^{\dagger} S^{-1} z)}{1 + (z_B^{\dagger} S_{BB}^{-1} z_B)} \geq \ell_0 \; . \tag{3-20}$$

Finally, if we define

$$v \equiv \frac{y}{[1 + (z_B^{\dagger} S_{BB}^{-1} z_B)]^{1/2}} \tag{3-21}$$

and substitute in Equation (3-17), we obtain the basic factorization formula

$$1 + (z^{\dagger} S^{-1} z) = [1 + (z_B^{\dagger} S_{BB}^{-1} z_B)] \left[1 + \frac{|v|^2}{T} \right] \; . \tag{3-22}$$

This relation also provides the final form for the LR test:

$$\ell = 1 + \frac{|v|^2}{T} \geq \ell_0 \; . \tag{3-23}$$

The statistical properties of T, y, and v introduced here are derived in the next section. It will then be found that the factorization given in Equation (3-22) has an

21

interesting interpretation which is fundamental to the evaluation of the performance of the likelihood ratio test.

4. PROBABILITY OF DETECTION OF THE LIKELIHOOD RATIO TEST

In order to evaluate the statistical performance of our decision rule, we begin by fixing the B components of all the data vectors. The conditional probability densities of y, v, and T will be obtained first, and the conditional probability of detection of the LR test will then be derived. Finally, the conditioning will be removed, by averaging over the joint PDF of the variables z_B and the $z_B(k)$. We remark at the outset that the conditional PDF of z_B will reflect the presence of that component of the actual signal which is orthogonal (after whitening) to the steering vector, and it is this feature which distinguishes the present analysis from that of Reference 1.

We begin with the quantity T, and substitute from the definition of S for the submatrices S_{AA}, S_{AB}, and S_{BA}:

$$T = S_{AA} - S_{AB} S_{BB}^{-1} S_{BA}$$

$$= \sum_{k=1}^{K} z_A(k) z_A(k)^* - \sum_{k,m=1}^{K} z_A(k) z_B(k)^\dagger S_{BB}^{-1} z_B(m) z_A(m)^* .$$

The K × K matrix

$$Q(k,m) \equiv z_B(k)^\dagger S_{BB}^{-1} z_B(m) \tag{4-1}$$

is constant under the conditioning, and we also define the matrix

$$R \equiv I_K - Q , \tag{4-2}$$

where I_K is the K × K identity. T is a quadratic form in the secondary A components, corresponding to the matrix R:

$$T = \sum_{k,m=1}^{K} z_A(k) R(k,m) z_A(m)^* . \tag{4-3}$$

The matrices Q and R are Hermitian, and we can reduce this quadratic form to a sum of squares by diagonalizing R. Let W be a unitary matrix which accomplishes this, so that we may write

$$R = W \mathcal{R} W^{\dagger} , \qquad\qquad (4\text{-}4)$$

where $\mathcal{R}$ is the diagonal matrix:

$$\mathcal{R} = \text{diag}\left[r_1, r_2, \ldots, r_K \right] .$$

The r_k are, of course, the eigenvalues of R.

When this representation is substituted for R, the quadratic form becomes

$$T = \sum_{m=1}^{K} r_m |w(m)|^2 , \qquad\qquad (4\text{-}5)$$

and the new random variables are given by

$$w(m) = \sum_{k=1}^{K} z_A(k) W(k,m) . \qquad\qquad (4\text{-}6)$$

The $z_A(k)$ are independent complex Gaussian scalars, with zero mean and unit variance, and the $w(k)$ have identical properties, under the conditioning, since they were obtained by means of a unitary transformation.

It can easily be verified by direct substitution that Q is idempotent and that its trace equals $N-1$ (see Reference 1 for details). It follows that R is also idempotent and that its trace equals

$$\text{Tr}[R] = K + 1 - N = L .$$

In this formula, L is the same parameter that was introduced in Equation (2-8).

Thus L of the r_k are equal to one, while the others vanish, and we can choose W so that the first L eigenvalues are unity. Then T is the simple sum

$$T = \sum_{m=1}^{L} |w(m)|^2 , \qquad\qquad (4\text{-}7)$$

which is a Chi-squared random variable with 2L degrees of freedom. With our convention, each complex random variable has unit variance, hence the conditional PDF of T is

$$f(T) = \frac{T^{L-1}}{(L-1)!} e^{-T} .$$ (4-8)

Since the actual values of the conditioning variables do not enter this formula for T, we see that Equation (4-8) gives the unconditioned PDF of T as well. This formula agrees with the known PDF of the maximum likelihood (Capon) spectral estimator.

Turning now to y, we have

$$y = z_A - S_{AB} S_{BB}^{-1} z_B$$

$$= z_A - \sum_{k=1}^{K} z_A(k) z_B(k)^\dagger S_{BB}^{-1} z_B .$$ (4-9)

Since y is a linear combination of all the A components, it is a complex scalar Gaussian random variable under the conditioning. In terms of the quantities

$$q(k) \equiv z_B(k)^\dagger S_{BB}^{-1} z_B ,$$ (4-10)

we have

$$y = z_A - \sum_{k=1}^{K} q(k) z_A(k) .$$ (4-11)

Since the secondaries have zero mean, the conditional expectation of y is just

$$E_B y = E z_A = b A_p \cos\theta e^{i\varphi} .$$

The A components are also independent with unit variance, hence the conditional variance of y is given by

$$E_B|y - E_B y|^2 = 1 + \sum_{k=1}^{K} |q(k)|^2$$

$$= 1 + (z_B^\dagger S_{BB}^{-1} z_B) . \tag{4-12}$$

The last step follows from direct substitution, using Equation (4-10).

As a result of this last evaluation, the variable v defined by Equation (3-21) has a conditional variance of unity:

$$E_B|v - E_B v|^2 = 1 ,$$

and its conditional mean can be written

$$E_B v = b A_p \cos\Theta e^{i\varphi} [1 + (z_B^\dagger S_{BB}^{-1} z_B)]^{-1/2} . \tag{4-13}$$

At this point we introduce the SNR loss factor ρ, by means of

$$\rho \equiv \frac{1}{1 + (z_B^\dagger S_{BB}^{-1} z_B)} , \tag{4-14}$$

and then we have the simple expression

$$|E_B v|^2 = |b|^2 A_p^2 \rho \cos^2\Theta \tag{4-15}$$

for the signal energy contained in v, under the conditioning.

It remains to be shown that y, and hence also v, is conditionally independent of T. We note that the $z_B(k)$, being K in number and of dimension $N-1$, must be linearly dependent (K must exceed $N-1$ in order to guarantee the non-singularity of the sample covariance matrices). By direct substitution we see that

$$\sum_{k=1}^{K} z_B(k) Q(k,m) = z_B(m) ,$$

and thus

$$\sum_{k=1}^{K} z_B(k)\, R(k,m) = 0 \; ,$$

which provides the correct number of linear relations, since the rank of R is L. From these relations it follows easily that

$$\sum_{k=1}^{K} R(m,k)\, z_B(k)^{\dagger} = 0$$

and hence that

$$\sum_{k=1}^{K} R(m,k)\, q(k) = 0 \; . \tag{4-16}$$

Equation (4-16) shows that the $q(k)$ are the components of a K vector which is an eigenvector of R, corresponding to the null eigenvalue. We can always arrange things, by a suitable choice of W, so that this vector is proportional to the (L+1)st eigenvector of R, in other words, to the first of the eigenvectors corresponding to the eigenvalue zero. We have already evaluated the norm of this vector, in deriving Equation (4-12). The proof of the independence of v and T is completed by inverting Equation (4-6) to express the $z_A(k)$ in terms of the $w(k)$, and substituting in Equation (4-11), with the result

$$y = z_A - \sum_{k,m=1}^{K} q(k)\, W(k,m)^{*}\, w(m) \; .$$

Recognizing that the columns of W are the eigenvectors of R, we finally obtain

$$y = z_A - (z_B^{\dagger} S_{BB}^{-1} z_B)^{1/2}\, w_{L+1} \; , \tag{4-17}$$

which shows that y (and thus v) is conditionally independent of the first L of the $w(k)$, and is therefore also conditionally independent of T.

Returning to the LR decision rule, given by Equation (3-23), we see that $\ell-1$ has a simple characterization, when conditioned on the B components of the data vectors:

$$\ell - 1 = \frac{|v|^2}{T} \; , \tag{4-18}$$

where T is Chi-squared and v is Gaussian, independent of T, with unit variance and (conditional) mean given by Equation (4-13). The PDF of this ratio is a particular case of the non-central F distribution, with two degrees of freedom for the (non-central) numerator, and 2L degrees of freedom for the (central) denominator. Expressed in another way, the conditional PD of this test, which we denote by $P_D(B)$, is identical to the PD of a simple CFAR decision rule for a radar which uses one hit (i.e., no noncoherent integration) and a threshold based on L complex samples of noise.

Although the non-central F distribution is well known,[4] the simple form which its cumulative probability distribution function assumes when the numbers of degrees of freedom are both even is less familiar. If we define the parameter

$$a \equiv |b|^2 A_p^2 \cos^2 \Theta \; , \tag{4-19}$$

then we have

$$|E_B v|^2 = a\rho \; . \tag{4-20}$$

The desired formula is derived in Section 5, and the result is

$$P_D(B) = 1 - \frac{1}{\ell_0^L} \sum_{k=1}^{L} \binom{L}{k} (\ell_0 - 1)^k G_k\left(\frac{a\rho}{\ell_0}\right) \; . \tag{4-21}$$

The function G which enters this expression is the incomplete Gamma function:

$$G_k(y) = \frac{1}{(k-1)!} \int_y^\infty e^{-u} u^{k-1} du$$

$$= e^{-y} \sum_{m=0}^{k-1} \frac{y^m}{m!} \; . \tag{4-22}$$

28

The conditioning variables survive in Equation (4-21) only in the loss factor, ρ. The PDF of this loss factor will also be derived in Section 5, which allows us to complete the evaluation of the PD of the decision rule. Denoting the PDF of ρ simply by $f(\rho)$, we obtain

$$P_D = 1 - \frac{1}{\ell_0^L} \sum_{k=1}^{L} \binom{L}{k} (\ell_0 - 1)^k H_k\left(\frac{a}{\ell_0}\right) , \qquad (4\text{-}23)$$

where

$$H_k(y) = \int_0^1 G_k(\rho y) f(\rho) \, d\rho . \qquad (4\text{-}24)$$

One of the forms of the PDF of the loss factor obtained in Section 5 yields the following formula for $f(\rho)$:

$$f(\rho) = \exp[-|b|^2 A_p^2 \sin^2 \Theta] \sum_{m=0}^{\infty} f_\beta(\rho; K+2-N, N-1+m) \frac{(|b|^2 A_p^2 \sin^2 \Theta)^m}{m!} . \qquad (4\text{-}25)$$

The integrals defined by Equation (4-24) are evaluated in the Appendix, which contains an expression for the detection probability suitable for numerical computation.

If the signal component of v is absent, then all dependence on the conditioning vanishes, hence we have already obtained the PFA of our test. Putting $a=0$ and noting that $G_k(0)=1$, we obtain from Equation (4-21) the simple formula

$$PFA = \frac{1}{\ell_0^L} . \qquad (4\text{-}26)$$

It is noteworthy that the signal-to-noise parameter, a, will vanish in our problem if $\Theta = \pi/2$, no matter how strong the actual signal may be. As characterized in Section 2, this is the case of a signal which falls in a null of the asymptotic adapted antenna pattern. Although the SNR itself (as defined in Section 2) can be large, our present result shows that the detection system responds as though this signal were entirely absent.

29

Returning to the factorization expressed by Equation (3-22), we observe that when $a = 0$ the two factors on the right side become completely independent. We introduce the notation

$$\Sigma(N, K) \equiv (z^{\dagger} S^{-1} z) , \qquad (4\text{-}27)$$

and use it only in the case in which each of the data vectors has a covariance matrix equal to the identity. The first argument of Σ refers to the dimensionality of the vectors, and the second to the number of secondaries, so there will be $K + 1$ independent Gaussian vectors in all. With this notation, we can write

$$(z_B^{\dagger} S_{BB}^{-1} z_B) = \Sigma(N - 1, K) .$$

Since v is a scalar, and since T/K can be interpreted as the one-dimensional sample covariance matrix of L "secondaries," we can even write

$$\frac{|v|^2}{T} = \Sigma(1, K + 1 - N) .$$

In terms of these new random variables, the factorization given in Equation (3-22), can be expressed in the form

$$1 + \Sigma(N, K) = [1 + \Sigma(N - 1, K)] [1 + \Sigma(1, K + 1 - N)] . \qquad (4\text{-}28)$$

When the means of the data vectors are zero, the two factors in square brackets on the right side of this identity are independent, and Equation (4-28) expresses a basic statistical property of the family of random variables $\Sigma(N,K)$.

Iterating Equation (4-28), and using the identity itself with $N = 2$, it can be shown that

$$1 + \Sigma(N, K) = [1 + \Sigma(N - 2, K)] [1 + \Sigma(2, K + 2 - N)] .$$

Continuing in this way, a more general factorization is obtained:

$$1 + \Sigma(N, K) = [1 + \Sigma(M, K)] [1 + \Sigma(N - M, K - M)] \qquad (4\text{-}29)$$

30

for any $M < N$. This last expression will be derived directly in Section 7, in connection with a generalization of the initial detection problem.

5. STATISTICAL PROPERTIES OF THE LOSS FACTOR

Consider a set of independent Gaussian vectors of dimension $\mathcal{N}$, consisting of a primary vector z and a set of secondaries z(k), where $1 \le k \le K$. The covariance matrix of each vector is the identity matrix $I_{\mathcal{N}}$, and the secondaries have zero mean. We assume that the primary vector has a signal component, given by

$$E z = \gamma t ,$$ (5-1)

where γ is a complex amplitude and t is a unit vector in the $\mathcal{N}$ space. Next, let

$$S \equiv \sum_{k=1}^{K} z(k) z(k)^{\dagger}$$

and

$$\Sigma(\mathcal{N}, K) \equiv (z^{\dagger} S^{-1} z) .$$

These quantities are just like those we have been discussing, but here we wish to analyze the statistical properties of $\Sigma(\mathcal{N}, K)$ and the related "loss factor,"

$$\rho(\mathcal{N}, K) \equiv \frac{1}{1 + \Sigma(\mathcal{N}, K)} ,$$ (5-2)

apart from the specific application in which they have arisen. The integer K will always represent the number of secondaries, which will not change, but $\mathcal{N}$ will be given different values in the several applications which are to be made of the results of this section.

The loss factor of Section 4, defined by Equation (4-14), will be statistically identical to $\rho(N-1,K)$, if we identify z with the component z_B of that section. The expected value of this primary component is given by Equation (3-9), and the identification of ρ with $\rho(N-1,K)$ will then be complete if we assign to the complex amplitude γ the value

$$\gamma = b A_p \sin \Theta$$ (5-3)

and associate the unit vector t with h.

33

If the signal component of the primary vector is absent, then $\Sigma(\mathcal{N},K)$ is the same as a quantity of this form discussed in Reference 1, where it was shown that in this case the corresponding loss factor is subject to the Beta distribution. In the general case we define

$$c \equiv |\gamma|^2 , \tag{5-4}$$

and we denote the PDF of $\rho(\mathcal{N},K)$ by $f(\rho;\mathcal{N},K,c)$, anticipating its independence of the unit vector t. Then, in the signal-free case, the known result is

$$f(\rho;\mathcal{N},K,0) = f_\beta(\rho;K+1-\mathcal{N},\mathcal{N}) . \tag{5-5}$$

Our notation for the Beta distribution was given in Equation (2-9).

One approach to the analysis of $\Sigma(\mathcal{N},K)$ would begin with a decomposition of all vectors into components parallel and orthogonal to the given unit vector t. A factorization just like that of Equation (4-28) would follow, and the PDF of $\Sigma(\mathcal{N}-1,K)$ is already known. The other, statistically independent factor encountered in this method can be shown to have a simple non-central F distribution, and the desired formula for $f(\rho;\mathcal{N},K,c)$ can then be obtained in a straightforward manner. We follow instead another route, which leads more directly to the objective, based on the suggestion of K. Forsythe[5] to condition first on the primary data vector z itself.

With z fixed, we introduce a unit vector d in the direction of z, and write

$$z = (z^\dagger z)^{1/2} d , \tag{5-6}$$

and then we have

$$\Sigma(\mathcal{N},K) = (z^\dagger z)(d^\dagger S^{-1} d) . \tag{5-7}$$

Next we carry out a rotation of the secondary vectors, by means of some unitary transformation, which will leave S unchanged statistically but will rotate d into the basis vector

$$e = [1,0,...,0]^T .$$

The inner product containing S on the right side of Equation (5-7) is identical to one studied in Section 3 and evaluated in Equation (3-18). It follows that

$$\Sigma(\mathcal{N},K) = \frac{(z^\dagger z)}{T} ,$$ (5-8)

where T is a Chi-squared random variable. We have not introduced a new notation for the rotated primary vector, since only its norm, which is invariant to such a transformation, appears in Equation (5-8).

The PDF of T is given by Equation (4-8), which we now write in the form

$$f(T) = \frac{T^{\mathcal{L}-1}}{(\mathcal{L}-1)!} e^{-T} ,$$

where

$$\mathcal{L} \equiv K+1-\mathcal{N} .$$ (5-9)

The integer $\mathcal{L}$ is directly analogous to L of the previous sections, but we use a separate symbol here, just as with $\mathcal{N}$, to avoid confusion when the results are applied to different cases.

As before, the statistical properties of T are independent of the values of the original conditioning variables, hence Equation (5-8) represents $\Sigma(\mathcal{N},K)$ as a ratio of independent random variables. The numerator is just the sum of the squares of $\mathcal{N}$ complex Gaussian variables, each of unit variance, hence its PDF is a non-central Chi-squared distribution. This latter PDF depends only on $\mathcal{N}$ and the sum of the squares of the expected values, which is the same as the squared norm

$$(Ez)^\dagger(Ez) = |\gamma|^2 = c .$$

We proceed with the analysis of the ratio of two Chi-squared random variables, making the definition

$$x(\mathcal{N},\mathcal{L}) = \frac{\displaystyle\sum_{n=1}^{\mathcal{N}} |z_n|^2}{\displaystyle\sum_{l=1}^{\mathcal{L}} |w_l|^2} ,$$ (5-10)

35

and treating $\mathcal{N}$ and $\mathcal{L}$ as independent variables. The z_n and the w_l are all complex Gaussian variables with unit variance; the w_l have zero mean, and we have, for the numerator,

$$\sum_{n=1}^{\mathcal{N}} |E z_n|^2 = c .$$

Our results will then apply to $\Sigma(\mathcal{N}, K)$ with the identification

$$\Sigma(\mathcal{N}, K) = x(\mathcal{N}, K+1-\mathcal{N}) . \qquad (5\text{-}11)$$

They will also provide the statistical properties of the likelihood ratio test of Section 4, according to Equation (4-18), since we can put

$$l = 1 + x(1, K+1-N)$$

and replace c by $a\rho$, in accordance with Equation (4-20). As noted, the PDF of the ratio $x(\mathcal{N}, \mathcal{L})$ is a special case of the non-central F distribution. The method to be used here, which results in an expression for the cumulative probability distribution function of this ratio in the form of a finite sum, was described in detail in Reference 6 which contains a number of similar computations.

Let us write

$$P_x(\xi; \mathcal{N}, \mathcal{L}) = \text{Prob}[x(\mathcal{N}, \mathcal{L}) \geq \xi] , \qquad (5\text{-}12)$$

so that the cumulative probability distribution function of the variable $x(\mathcal{N}, \mathcal{L})$ is $1 - P_x(\xi; \mathcal{N}, \mathcal{L})$. With $x(\mathcal{N}, \mathcal{L})$ we associate the random variable $y(\mathcal{N}, \mathcal{L})$:

$$y(\mathcal{N}, \mathcal{L}) \equiv \frac{1}{1 + x(\mathcal{N}, \mathcal{L})} , \qquad (5\text{-}13)$$

so that, in analogy to Equation (5-11), we have

$$\rho(\mathcal{N}, K) = y(\mathcal{N}, K+1-\mathcal{N}) . \qquad (5\text{-}14)$$

While $x(\mathcal{N}, \mathcal{L})$ ranges over all non-negative values, $y(\mathcal{N}, \mathcal{L})$ is confined to the interval between zero and one. The cumulative probability distribution function of $y(\mathcal{N}, \mathcal{L})$ will be denoted by

$$F_y(r;\mathcal{N},\mathcal{L}) = \text{Prob}[y(\mathcal{N},\mathcal{L}) \leq r] = \text{Prob}[x(\mathcal{N},\mathcal{L}) \geq \frac{1-r}{r}] = P_x(\frac{1-r}{r};\mathcal{N},\mathcal{L}) .$$

The PDF of $y(\mathcal{N},\mathcal{L})$, which will give us the desired PDF of the loss factor, will then be

$$f_y(r;\mathcal{N},\mathcal{L}) = \frac{d}{dr} F_y(r;\mathcal{N},\mathcal{L}) .$$

If we let

$$\eta \equiv \sum_{n=1}^{\mathcal{N}} |z_n|^2 - \xi \sum_{l=1}^{\mathcal{L}} |w_l|^2 , \tag{5-15}$$

then

$$P_x(\xi;\mathcal{N},\mathcal{L}) = \text{Prob}(\eta \geq 0) .$$

If, further, we denote the characteristic function of η by

$$\Phi(\lambda) \equiv E\,e^{i\lambda\eta} ,$$

the PDF of η will be

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi(\lambda)\,e^{-i\lambda\eta}\,d\lambda ,$$

and therefore

$$P_x(\xi;\mathcal{N},\mathcal{L}) = \frac{1}{2\pi} \int_0^\infty \left[\int_{-\infty}^{\infty} \Phi(\lambda)\,e^{-i\lambda\eta}\,d\lambda \right] d\eta . \tag{5-16}$$

The contour of integration in the λ plane can be moved an infinitesimal distance ε below the real axis, without changing the value of the integral. This is so because, as will be seen shortly, the characteristic function is analytic in a finite band containing

the real axis. After this change, the order of integration may be reversed in Equation (5-16), and the integral over η carried out explicitly. Since the imaginary part of λ is negative, we obtain the simple result

$$P_x(\xi;\mathcal{N},\mathcal{L}) = \frac{1}{2\pi i} \int_{-\infty-i\varepsilon}^{\infty-i\varepsilon} \Phi(\lambda)\,\frac{d\lambda}{\lambda} \ . \tag{5-17}$$

This basic technique has been used by a number of authors to obtain the cumulative probability distribution function of the ratio of two independent random variables.

The characteristic functions of the two Chi-squared variables encountered in η are well known, and the characteristic function of η is easily evaluated as the product

$$\Phi(\lambda) = e^{-c}\,\frac{e^{c/(1-i\lambda)}}{(1-i\lambda)^{\mathcal{N}}}\,\frac{1}{(1+i\xi\lambda)^{\mathcal{L}}} \ .$$

This function has a pole of order $\mathcal{L}$ at $\lambda = i/\xi$, and an essential singularity at $\lambda = -i$. It is analytic elsewhere, including, of course, a finite band containing the real axis.

We substitute in Equation (5-17), complete the contour at infinity below the real axis (the integrand vanishes rapidly as $|\lambda| \to \infty$), and change the sense of traversal of the contour to positive. This yields

$$P_x(\xi;\mathcal{N},\mathcal{L}) = -\ \frac{1}{2\pi i}e^{-c}\oint_{|\lambda+i|=\varepsilon}e^{c/(1-i\lambda)}\,\frac{d\lambda}{\lambda\,(1-i\lambda)^{\mathcal{N}}\,(1+i\xi\lambda)^{\mathcal{L}}} \ . \tag{5-18}$$

The essential singularity is now shifted to the origin by the change of variable

$$\lambda = -i\,(1-t) \ ,$$

which results in the formula

$$P_x(\xi;\mathcal{N},\mathcal{L}) = \frac{1}{2\pi i}e^{-c}\oint_{|t|=\varepsilon}e^{c/t}\,\frac{dt}{(1-t)\,t^{\mathcal{N}}\,(1+\xi-\xi t)^{\mathcal{L}}} \ .$$

Next, we replace t by 1/t, so that

$$P_x(\xi; \mathcal{N}, \mathcal{L}) = \frac{1}{2\pi i} e^{-c} \oint_{|t|=1/\varepsilon} e^{ct} \frac{t^{\mathcal{N}+\mathcal{L}-1} dt}{(t-1)\left[(1+\xi)t - \xi\right]^{\mathcal{L}}} ,$$

and the path of integration is now a large circle of radius $1/\varepsilon$. This contour is now shrunk and broken into two small circles: one encircling the simple pole at $t = 1$, the other encircling the pole at

$$t_0 \equiv \frac{\xi}{1+\xi} .$$

The residue at $t = 1$ is unity, as can be seen by inspection, and hence

$$P_x(\xi; \mathcal{N}, \mathcal{L}) = 1 - \frac{1}{2\pi i} e^{-c} \oint_{|t-t_0|=\varepsilon} e^{ct} \frac{t^{\mathcal{N}+\mathcal{L}-1} dt}{(1-t)\left[(1+\xi)t - \xi\right]^{\mathcal{L}}} .$$

With a final change of variable, the remaining singularity is shifted to the origin in the new variable, u:

$$t \equiv \frac{\xi + u}{1 + \xi} ,$$

and the integral becomes

$$P_x(\xi; \mathcal{N}, \mathcal{L}) = 1 - \frac{e^{-c/(1+\xi)}}{(1+\xi)^{\mathcal{N}+\mathcal{L}-1}} \frac{1}{2\pi i} \oint_{|u|=\varepsilon} e^{cu/(1+\xi)} (\xi + u)^{\mathcal{N}+\mathcal{L}-1} \frac{du}{u^{\mathcal{L}}(1-u)} . \qquad (5\text{-}19)$$

Before proceeding with this derivation, we make the substitution

$$\xi = \frac{1-r}{r} \qquad (5\text{-}20)$$

in Equation (5-19), in order to get the cumulative probability distribution function of the random variable y:

$$F_y(r;\mathcal{N},\mathcal{L}) = 1 - \frac{1}{2\pi i} \oint_{|u|=\varepsilon} e^{-cr(1-u)} \left[1-r(1-u)\right]^{\mathcal{N}+\mathcal{L}-1} \frac{du}{u^{\mathcal{L}}(1-u)} \; .$$

The PDF of y is now obtained by differentiation under the integral sign:

$$f_y(r;\mathcal{N},\mathcal{L}) = \frac{1}{2\pi i} e^{-cr}$$

$$\times \oint_{|u|=\varepsilon} e^{cru} \left[c\,(1-r+ru)^{\mathcal{N}+\mathcal{L}-1} + (\mathcal{N}+\mathcal{L}-1)(1-r+ru)^{\mathcal{N}+\mathcal{L}-2} \right] \frac{du}{u^{\mathcal{L}}} \; . \qquad (5\text{-}21)$$

It proves to be much simpler to develop $f_y(r;\mathcal{N},\mathcal{L})$ from this integral than to complete the evaluation of $F_y(r;\mathcal{N},\mathcal{L})$ and then differentiate the result.

Returning to Equation (5-19), we make the binomial expansion

$$(\xi+u)^{\mathcal{N}+\mathcal{L}-1} = \sum_{m=0}^{\mathcal{N}+\mathcal{L}-1} \binom{\mathcal{N}+\mathcal{L}-1}{m} \xi^m \, u^{\mathcal{N}+\mathcal{L}-1-m}$$

and substitute in the integral. Only for the terms with negative resultant powers of u are there poles in the integrand, and we have then

$$P_x(\xi;\mathcal{N},\mathcal{L}) = 1 - \frac{e^{-c/(1+\xi)}}{(1+\xi)^{\mathcal{N}+\mathcal{L}-1}}$$

$$\times \sum_{m=\mathcal{N}}^{\mathcal{N}+\mathcal{L}-1} \binom{\mathcal{N}+\mathcal{L}-1}{m} \xi^m \, \frac{1}{2\pi i} \oint_{|u|=\varepsilon} e^{cu/(1+\xi)} \frac{du}{u^{m+1-\mathcal{N}}(1-u)} \; . \qquad (5\text{-}22)$$

The contour integral

$$G_n(y) \equiv e^{-y} \frac{1}{2\pi i} \oint_{|u|=\varepsilon} e^{yu} \frac{du}{u^n(1-u)}$$

is easily evaluated by expanding the factor (1−u) in the denominator:

$$G_n(y) = e^{-y} \sum_{k=1}^{n} \frac{1}{2\pi i} \oint_{|u|=\varepsilon} e^{yu} \frac{du}{u^k}$$

$$= e^{-y} \sum_{k=0}^{n-1} \frac{y^k}{k!} . \tag{5-23}$$

This expression is exactly the form of the incomplete Gamma function which was defined earlier [see Equation (4-22)]. We substitute in Equation (5-22) and, after a redefinition of the summation index, obtain the final result

$$P_x(\xi;\mathcal{N},\mathcal{L}) = 1 - \frac{\xi^{\mathcal{N}}}{(1+\xi)^{\mathcal{N}+\mathcal{L}-1}} \sum_{m=0}^{\mathcal{L}-1} \binom{\mathcal{N}+\mathcal{L}-1}{\mathcal{N}+m} \xi^m G_{m+1}\left(\frac{c}{1+\xi}\right) . \tag{5-24}$$

As noted earlier, the detection probability of the LR test is a special case which corresponds to the parameter assignments

$$\mathcal{N} = 1$$
$$\mathcal{L} = K+1-N = L$$
$$c = a\rho . \tag{5-25}$$

According to Equations (4-18) and (3-23), the appropriate threshold is $\xi = \ell_0 - 1$. When these substitutions are made, Equation (5-24) becomes identical to Equation (4-21), which completes our derivation of that result.

It remains to obtain the PDF of the general loss factor, and hence we step back to Equation (5-21). Making a binomial expansion as before, we evaluate the integral

$$\frac{1}{2\pi i} \oint_{|u|=\varepsilon} e^{cru} (1-r+ru)^{\mathcal{N}+\mathcal{L}-1} \frac{du}{u^{\mathcal{L}}}$$

$$= r^{\mathcal{L}-1}(1-r)^{\mathcal{N}} \sum_{m=0}^{\mathcal{L}-1} \binom{\mathcal{N}+\mathcal{L}-1}{\mathcal{N}+m} \frac{(1-r)^m c^m}{m!} .$$

The other term in Equation (5-21) involves the same integral, but with $\mathcal{N}$ replaced by $\mathcal{N}-1$. When the results are combined we find

$$f_y(r;\mathcal{N},\mathcal{L}) = e^{-cr}\, r^{\mathcal{L}-1}(1-r)^{\mathcal{N}-1}$$

$$\times \sum_{m=0}^{\mathcal{L}-1}\left[\binom{\mathcal{N}+\mathcal{L}-1}{\mathcal{N}+m}\frac{(1-r)^{m+1}c^{m+1}}{m!} + (\mathcal{N}+\mathcal{L}-1)\binom{\mathcal{N}+\mathcal{L}-2}{\mathcal{N}+m-1}\frac{(1-r)^m c^m}{m!}\right].$$

The summation in the second line of this equation can be written

$$\frac{(\mathcal{N}+\mathcal{L}-1)!}{(\mathcal{L}-1)!}\sum_{m=0}^{\mathcal{L}-1}\binom{\mathcal{L}-1}{m}\left[\frac{(1-r)^{m+1}c^{m+1}}{(\mathcal{N}+m)!} + \frac{(1-r)^m c^m}{(\mathcal{N}+m-1)!}\right]$$

$$= \frac{(\mathcal{N}+\mathcal{L}-1)!}{(\mathcal{L}-1)!}\sum_{m=0}^{\mathcal{L}}\left[\binom{\mathcal{L}-1}{m-1} + \binom{\mathcal{L}-1}{m}\right]\frac{(1-r)^m c^m}{(\mathcal{N}+m-1)!}.$$

The vanishing of the binomial coefficients (when the lower member is either negative or exceeds the upper member) controls the limits of summation, and with a standard binomial identity

$$\binom{n-1}{m-1} + \binom{n-1}{m} = \binom{n}{m} \tag{5-26}$$

we obtain the result:

$$f_y(r;\mathcal{N},\mathcal{L}) = e^{-cr}\, r^{\mathcal{L}-1}(1-r)^{\mathcal{N}-1}\frac{(\mathcal{N}+\mathcal{L}-1)!}{(\mathcal{L}-1)!}\sum_{m=0}^{\mathcal{L}}\binom{\mathcal{L}}{m}\frac{(1-r)^m c^m}{(\mathcal{N}+m-1)!}. \tag{5-27}$$

This formula can be expressed in several interesting forms. In terms of the Beta distribution, we have

$$f_y(r;\mathcal{N},\mathcal{L}) = e^{-cr}\sum_{m=0}^{\mathcal{L}}\binom{\mathcal{L}}{m}\frac{(\mathcal{N}+\mathcal{L}-1)!}{(\mathcal{N}+\mathcal{L}-1+m)!}\, c^m\, f_\beta(r;\mathcal{L},\mathcal{N}+m). \tag{5-28}$$

Another useful form is

$$f_y(r;\mathcal{N},\mathcal{L}) = f_\beta(r;\mathcal{L},\mathcal{N}) \, e^{-cr} \sum_{m=0}^{\mathcal{L}} \frac{\mathcal{L}! \, (\mathcal{N}-1)!}{(\mathcal{L}-m)! \, (\mathcal{N}-1+m)!} \frac{(1-r)^m \, c^m}{m!}$$

$$= f_\beta(r;\mathcal{L},\mathcal{N}) \, e^{-cr} \, {}_1F_1[-\mathcal{L};\mathcal{N};-(1-r)c] , \qquad (5\text{-}29)$$

where ${}_1F_1$ is the confluent hypergeometric function[7]. In the present case, the series for this function is finite.

An application of Kummer's first transformation (see Reference 7) yields the expression

$$f_y(r;\mathcal{N},\mathcal{L}) = f_\beta(r;\mathcal{L},\mathcal{N}) \, e^{-c} \, {}_1F_1[\mathcal{N}+\mathcal{L};\mathcal{N};(1-r)c] , \qquad (5\text{-}30)$$

in which the r-dependence has been removed from the exponential. If the infinite series is now substituted for the hypergeometric function, this formula assumes the interesting form

$$f_y(r;\mathcal{N},\mathcal{L}) = e^{-c} \sum_{m=0}^{\infty} f_\beta(r;\mathcal{L},\mathcal{N}+m) \, \frac{c^m}{m!} . \qquad (5\text{-}31)$$

The fact that this PDF is properly normalized is now evident by inspection, since each Beta PDF is so normalized. It is clear from either of these formulas that in the absence of a signal component we have simply

$$f_y(r;\mathcal{N},\mathcal{L}) = f_\beta(r;\mathcal{L},\mathcal{N}) ,$$

which is consistent with Equation (5-5), with $\mathcal{L} = K+1-\mathcal{N}$.

The mean value of the random variable y can be obtained from Equation (5-31), by using the formula

$$\int_0^1 f_\beta(x;n,m) \, x \, dx = \frac{n}{n+m}$$

for the mean of the Beta distribution. This yields the series

$$Ey = e^{-c} \sum_{m=0}^{\infty} \frac{\mathcal{L}}{\mathcal{N}+\mathcal{L}+m} \frac{c^m}{m!}$$

$$= \frac{\mathcal{L}}{\mathcal{N}+\mathcal{L}} e^{-c} \, {}_1F_1[\mathcal{N}+\mathcal{L};\mathcal{N}+\mathcal{L}+1;c]$$

for the expected value of y. Using the Kummer transformation again, this expression becomes

$$Ey = \frac{\mathcal{L}}{\mathcal{N}+\mathcal{L}} \, {}_1F_1[1;\mathcal{N}+\mathcal{L}+1;-c] \, ,$$

which can be expressed in the form

$$Ey = \mathcal{L} \frac{(\mathcal{N}+\mathcal{L}-1)!}{c^{\mathcal{N}+\mathcal{L}}} (-1)^{\mathcal{N}+\mathcal{L}} \left[e^{-c} - \sum_{m=0}^{\mathcal{N}+\mathcal{L}-1} \frac{(-c)^m}{m!} \right] . \tag{5-32}$$

When the signal component becomes large this PDF becomes more and more concentrated toward small values, and the asymptotic value of its mean, as obtained from Equation (5-32), is simply

$$Ey \xrightarrow[c\to\infty]{} \frac{\mathcal{L}}{c} . \tag{5-33}$$

According to the correspondence established in Equation (5-14), the PDF of the general loss factor, $\rho(\mathcal{N},K)$, is given by

$$f(\rho;\mathcal{N},K,c) = f_y(\rho;\mathcal{N},K+1-\mathcal{N})$$

$$= f_\beta(\rho;K+1-\mathcal{N},\mathcal{N}) \, e^{-c\rho} \, {}_1F_1[\mathcal{N}-K-1;\mathcal{N};-(1-\rho)c]$$

$$= e^{-c} \sum_{m=0}^{\infty} f_\beta(\rho;K+1-\mathcal{N},\mathcal{N}+m) \frac{c^m}{m!} . \tag{5-34}$$

The PDF of the loss factor which we require to complete the derivation of the performance of the LR decision rule is obtained by making the following parameter assignments in our result:

44

$$\mathcal{N} = N - 1$$

$$c = |b|^2 A_p^2 \sin^2 \Theta \ . \tag{5-35}$$

This loss factor was defined in Equation (4-14), and the signal component was given in Equation (3-9). When the parameters specified by Equation (5-35) are substituted in Equation (5-34), the formula quoted in Equation (4-25) is obtained.

6. MISMATCHED SIGNALS; NUMERICAL RESULTS AND DISCUSSION

The problem formulated in Section 3 involves the detection of a signal hypothesized to have the direction of a unit vector q in the N dimensional observation space. The probability of detection was then sought for the mismatched case, in which the primary sample contains a signal of amplitude b with a direction corresponding to a different unit vector p. The covariance matrix of the noise, originally unknown, is assumed to be M.

The quantity

$$SNR_{pp} \equiv |b|^2 (p^\dagger M^{-1} p) = |b|^2 A_p^2 , \qquad (6\text{-}1)$$

which we call the "available SNR" of this signal, represents the maximum SNR which could be attained if the noise covariance matrix were known *a priori* and if the system were steered for this signal direction. In fact, we are steering for q, and we are predominantly interested to know how rapidly the detection probability is reduced as the directions of p and q diverge. This will depend on the noise structure, represented by M, and also on the quality of our estimate of M, which is related to the number of secondary vectors K.

What we have found from our analysis is that the available SNR is split into two components, as follows:

$$SNR_{det} \equiv SNR_{pp} \cos^2 \Theta \qquad (6\text{-}2)$$

and

$$SNR_{loss} \equiv SNR_{pp} \sin^2 \Theta , \qquad (6\text{-}3)$$

where Θ is a measure of the separation of the p and q directions, relative to the matrix M:

$$\cos^2 \Theta \equiv \frac{|(q^\dagger M^{-1} p)|^2}{(q^\dagger M^{-1} q)(p^\dagger M^{-1} p)} . \qquad (6\text{-}4)$$

This angle Θ is zero only if the p and q directions coincide (since M is non-singular), and the relation of Θ to the asymptotic adapted antenna gain was discussed in Section 2.

The performance of our detection algorithm is identical to that of a simple CFAR detector which uses one sample for detection and L samples for threshold determination ($L = K + 1 - N$). The signal in the former sample has a signal to noise ratio equal to

$$SNR = SNR_{det} \; \rho \; , \tag{6-5}$$

where ρ is a random loss factor lying between zero and unity in value. This is analogous to the fluctuating signal models often used in radar analysis, but the loss factor which enters here has a more complicated PDF, which we have denoted by

$$f(\rho; N-1, K, SNR_{loss}) \; .$$

Thus part of the available SNR plays the role of a conventional signal, while the remainder affects the PDF of the loss factor ρ. The result is the effective SNR given by Equation (6-5).

Without mismatch, $SNR_{loss} = 0$ and the PDF of ρ reduces to a simple Beta distribution. The effect of this loss factor, together with the threshold estimation feature which gives the algorithm its CFAR character, was discussed in detail in Reference 1. If the number of secondary samples is sufficient, say $K = 5N$, then the total performance degradation in this case can be kept below 2 dB, relative to a system in which the noise background is perfectly known.

For mismatched signals, the effect of the component SNR_{loss} on the loss factor is to shift its PDF toward smaller values. Therefore, as Θ increases, the probability of detection will fall due to the combined effects of a smaller SNR_{det} as well as a greater loss due to the reduced values of ρ. As we have seen, the expected value of ρ is given by Equation (5-33), which can be written

$$\bar{\rho} \; \rightarrow \; \frac{K + 2 - N}{SNR_{loss}} \; ,$$

when SNR_{loss} itself is large. The average value of the effective SNR therefore tends to a constant as the signal amplitude increases for a fixed value of Θ:

$$\overline{SNR} = SNR_{det} \; \bar{\rho} \; \rightarrow \; (K + 2 - N) \cot^2 \Theta \tag{6-6}$$

48

as SNR_{pp} tends to infinity. We also note that $\text{SNR}_{det} = 0$ when $\Theta = \pi/2$, regardless of signal amplitude. This is the case in which a signal falls in a true null of the asymptotic adapted antenna pattern, as discussed in Section 2.

Using the results of Section 5, we can study the effect of the parameter SNR_{loss} on the loss factor probability density function. Figure 6-1 shows a series of curves of this PDF for the values $N = 4$ and $K = 20$, and for several values of SNR_{loss}, described on the legend as "signal component." Case A represents a matched signal, and the curve is a simple Beta distribution which peaks around 0.9. Signal components much less than 0 dB have little effect on the loss factor, but large values of SNR_{loss} shift its PDF dramatically toward small values.

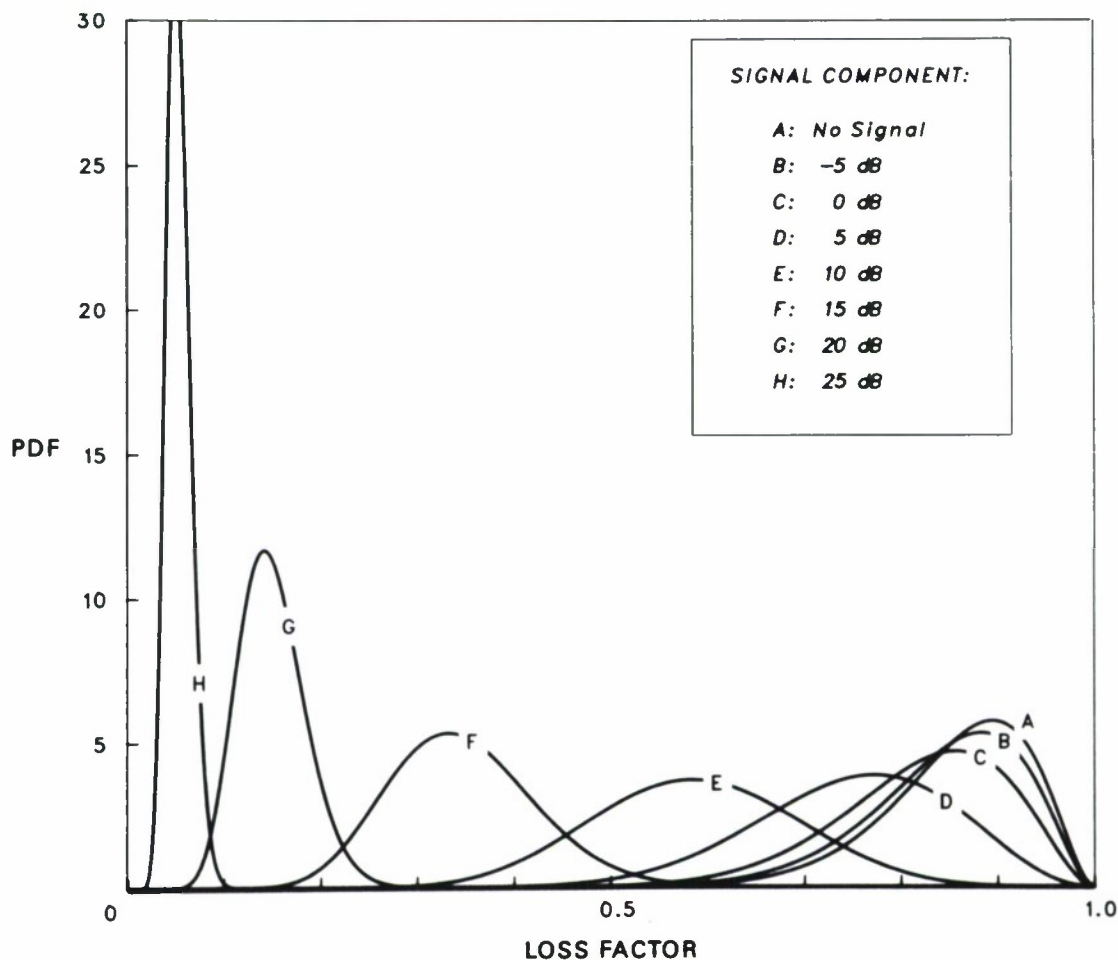Figure 6-1. Loss factor probability density function; N = 4 and K = 20.

49

Figure 6-2 is analogous, but it corresponds to an observation space of dimension 20 instead of 4. The number of secondaries has been increased in direct proportion. The signal-free case is again a simple Beta distribution, not quite so broad as before, and peaking at a loss factor of about 0.9 dB. The effect of a mismatched signal is much the same, but somewhat larger values of SNR_{loss} are required to achieve corresponding shifts of the loss factor PDF, relative to the parameters of Figure 6-1.

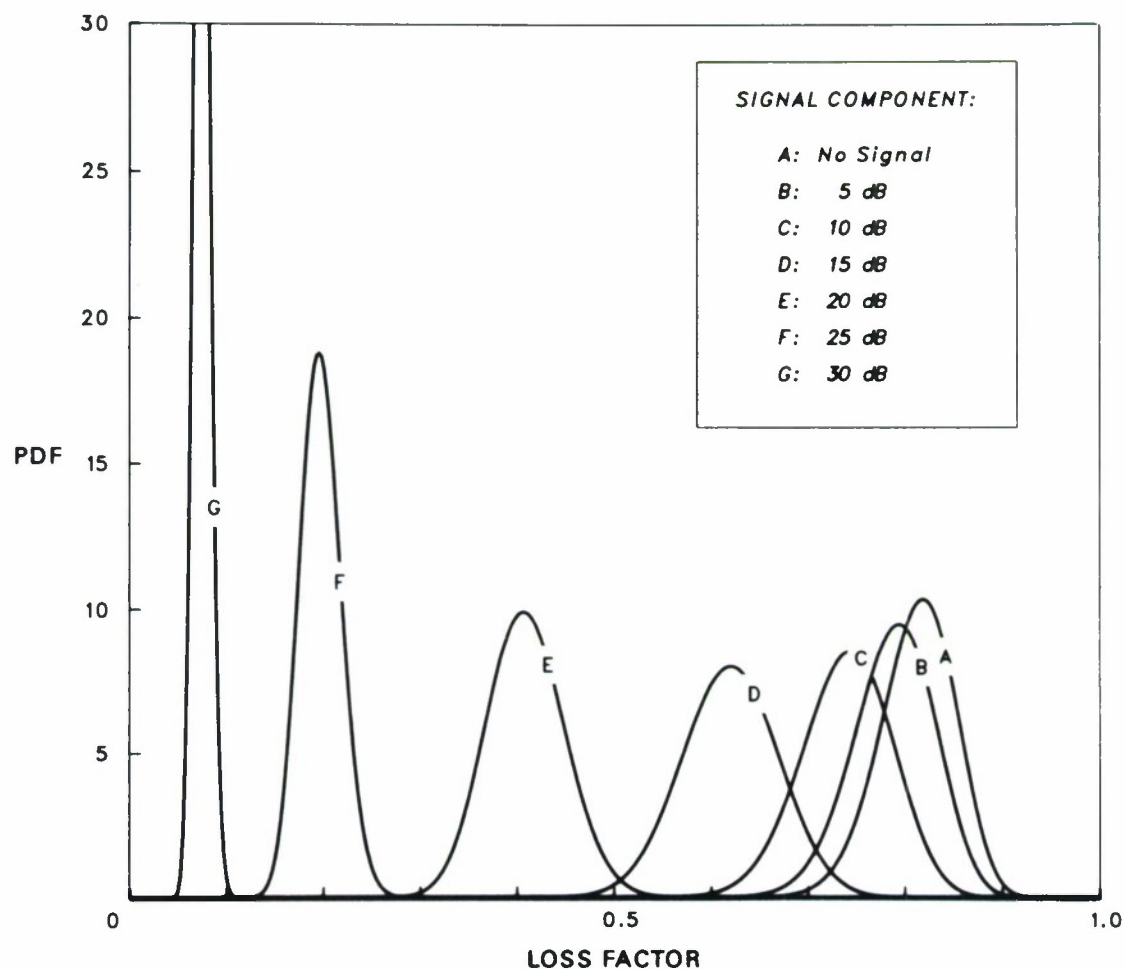*Figure 6-2. Loss factor probability density function; N = 20 and K = 100.*

The explicit expression for the probability of detection of a mismatched signal is fairly complicated, and a full derivation is presented in the Appendix. The basic parameters are the assigned PFA, the signal strength and the direction of the arriving signal relative to the steering direction. The value $PFA = 10^{-6}$ has been used for all

the numerical results given here. It is a representative number and, in general, system performance does not change dramatically as a function of PFA. Signal strength is parametrized by the "available SNR" (SNR_{pp}) defined in Equation (6-1). This parameter represents the strength of a signal relative to the actual interference present, and hence a constant value of SNR_{pp} will usually imply a varying signal amplitude, as the direction of the signal in the observation space changes.

The direction of the signal relative to the steering vector is parametrized by the quantity $\cos^2 \Theta$, defined in Equation (6-4). This is the most important parameter for our discussion, and its significance should be clearly understood. According to Equation (2-4), this is the factor by which the available SNR of a signal is reduced when its direction differs from that of the steering vector by the angle Θ in the observation space. If the interference is only white noise, then this factor is identical to the side-lobe gain, relative to the peak gain, of the antenna system in the signal direction. In the presence of directive interference, it describes the factor by which a signal of constant available SNR is rejected by the asymptotic adapted pattern, even though the adaptation was carried out with respect to (known) interference which does not include this unwanted signal. The detection performance results obtained in this study then show how such a signal is rejected when the interference is not known, but is estimated from secondary data from which this particular signal is absent.

A typical set of curves of detection probability versus SNR is shown in Figure 6-3, with fixed PFA and with $\cos^2 \Theta$ as the parameter for each curve. Curve A describes the matched case, $\Theta = 0$, while the others show the rapidly decreasing detection probability as $\cos^2 \Theta$ decreases. The smallest value shown corresponds to a reduction of only slightly more than 5 dB in effective sidelobe gain, yet the detection probability is sharply reduced. If $\cos^2 \Theta$ were zero, then the detection probability curve would be flat at the value PFA, as already noted. It should be kept in mind, however, that there may be no true nulls in the asymptotic adapted pattern; in other words, the value $\Theta = \pi/2$ may be physically unrealizable.

With increasing SNR, the curves of detection probability eventually level off at asymptotic values between the PFA and unity. We have noted that the mean value of the effective SNR [see Equation (6-6)] tends to a constant as the actual signal strength tends to infinity, and it can be shown that the standard deviation of this quantity also tends to a constant, non-zero value. In the Appendix an expression is given for the limiting form of the detection probability, for infinite signal amplitude, as a function of $\cos^2 \Theta$ and the other parameters of the problem. This limiting probability does not depend on N and K separately, but only on the parameter $L = K + 1 - N$, which measures the excess of the number of secondaries over the minimum necessary for covariance matrix estimation.
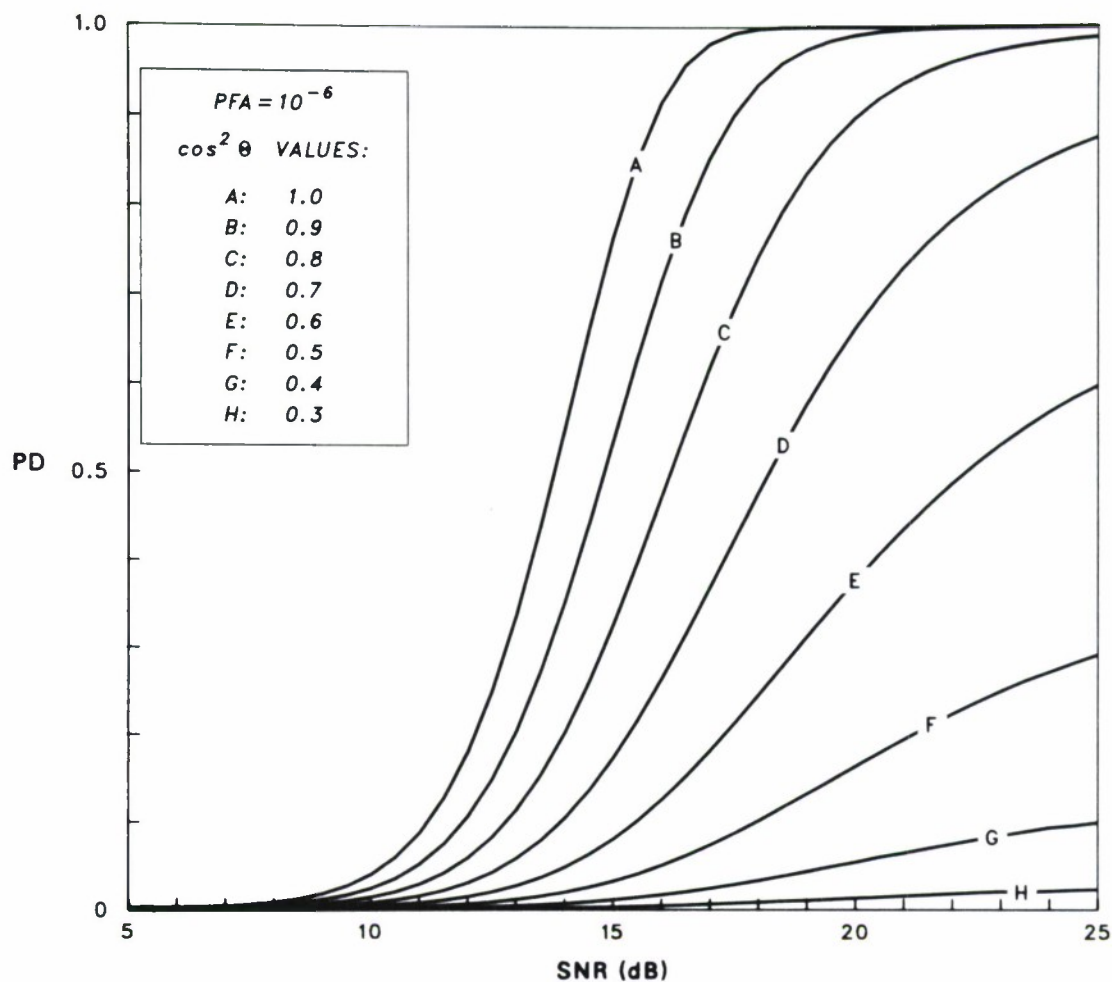
Figure 6-3. Mismatched signal probability of detection; N = 4 and K = 20.

A plot of this asymptotic (in the sense of large signal amplitude) detection probability is given in Figure 6-4. The independent parameter is $\sin^2\theta$, and each curve corresponds to a different value of L. The most striking feature of these curves is their tendency to crowd toward the right as the number of excess secondaries gets larger. This could have been anticipated from the form of Equation (6-6), which shows that the mean effective SNR increases linearly with L. When L itself is very large, we are approaching the case of known interference, since the quality of the estimate of its covariance continues to improve. But when the interference is known, the appropriate detector is a simple colored noise matched filter, and as long as $\cos^2\theta$ is non-zero, the SNR developed by any signal eventually tends to infinity. In other words, if a signal does not fall into a true null of the adapted pattern, it will ultimately
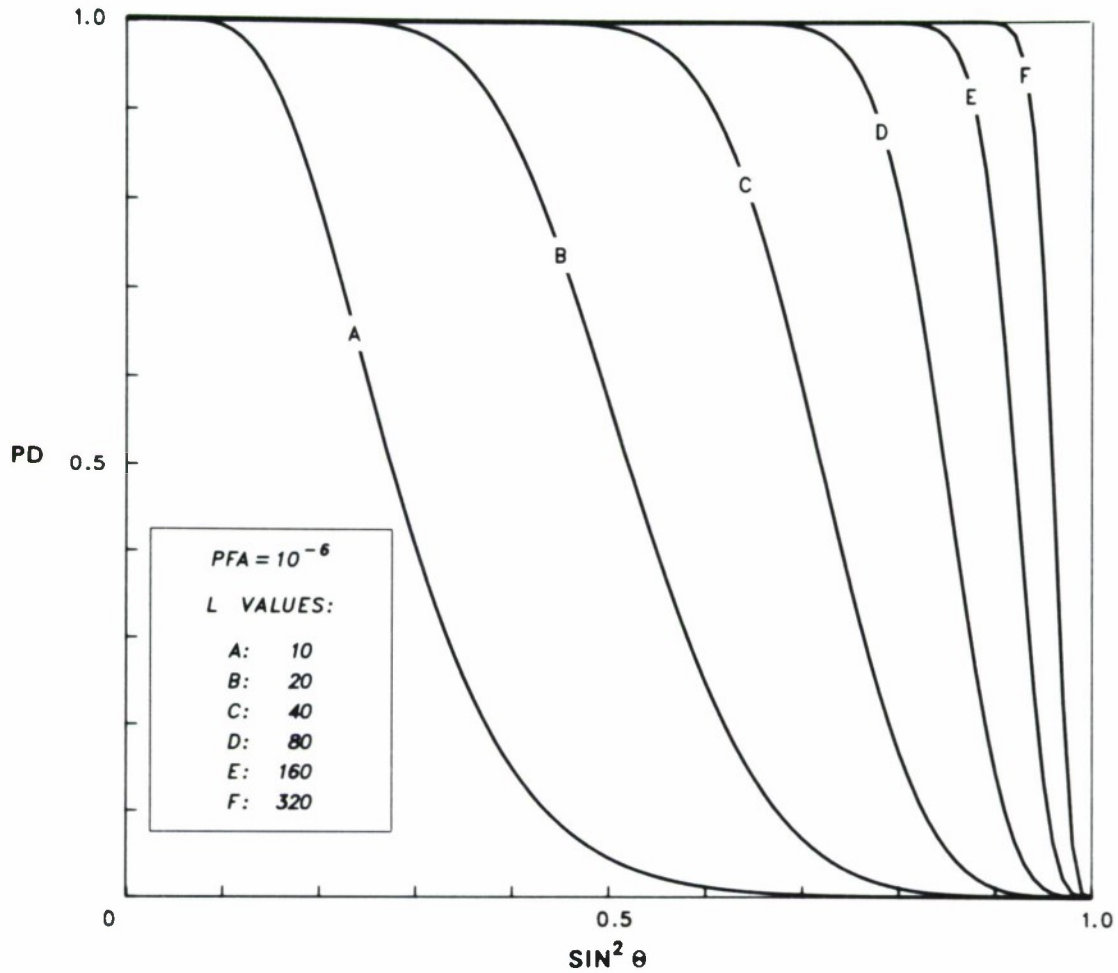
Figure 6-4. Asymptotic probability of detection.

burn through the sidelobes as its amplitude increases without bound, as indicated by Equation (2-4).

When the interference is unknown, however, the LR decision rule itself involves both interference estimation and threshold setting, and a true CFAR detector results. In effect, the component of an arriving signal which is orthogonal to the direction of the desired signal is included in the threshold estimation data, lowering the ultimate threshold and reducing the effect of this undesired signal. This mechanism can be appreciated by recasting the decision rule in the following way.

The LR decision rule, given in Equation (3-1), involves the matrix S which is K times the sample covariance matrix of the secondary data vectors. We can therefore write

$$KS^{-1} = \widehat{M}^{-1} ,$$

where $\widehat{M}$ is the sample covariance, an estimator of M. The LR test can then be expressed in terms of $\widehat{M}$ as follows:

$$\frac{K + (z^\dagger \widehat{M}^{-1} z)}{K + (z^\dagger \widehat{M}^{-1} z) - \dfrac{|(q^\dagger \widehat{M}^{-1} z)|^2}{(q^\dagger \widehat{M}^{-1} q)}} \geq \ell_0 . \tag{6-7}$$

We introduce the notation [a,b] for the inner product of two vectors, relative to the inverse of the sample covariance matrix:

$$[a, b] \equiv (a^\dagger \widehat{M}^{-1} b) .$$

Then the LR decision rule can be expressed as

$$\frac{K + [z, z]}{K + [z, z] - \dfrac{|[q, z]|^2}{[q, q]}} \geq \ell_0 . \tag{6-8}$$

Of course, [z,z] and [q,q] are the norms of the vectors z and q, associated with this inner product.

Next, we introduce the unit vector corresponding to the steering vector, q:

$$e_q \equiv [q, q]^{-1/2} q ,$$

and the components

$$z_\| \equiv [e_q, z] e_q$$

and

$$z_\perp \equiv z - z_\|$$

54

of z, which are respectively parallel and orthogonal to q in the sense of this inner product. It is easily verified that

$$[z_\perp, z_\parallel] = 0$$

and

$$[z, z] = [z_\parallel, z_\parallel] + [z_\perp, z_\perp] \, ,$$

and also that

$$\frac{\|[q, z]\|^2}{[q, q]} = \|[e_q, z]\|^2 = [z_\parallel, z_\parallel] \, .$$

With these definitions we obtain the formula

$$\frac{K + [z, z]}{K + [z_\perp, z_\perp]} \geq \ell_0$$

for the LR test, which in turn can be simplified to

$$\frac{[z_\parallel, z_\parallel]}{1 + \frac{1}{K}[z_\perp, z_\perp]} \geq K(\ell_0 - 1) \, . \tag{6-9}$$

This formula shows clearly that any component of z which is orthogonal to the steering vector q, in the sense of this inner product (which depends on the *estimate* of the interference matrix), will serve to raise the effective threshold for detection. The formula also shows how this feature disappears in the limit of large K. Finally, in this limit we see that the LR test becomes the familiar colored noise matched filter detector, since $\hat{M}$ converges to M when K increases without bound.

7. A GENERALIZATION OF THE DETECTION PROBLEM

The problem we have been discussing so far concerns the detection of a signal of unknown complex amplitude and known direction in an N dimensional space of data vectors. In addition, the covariance matrix of the background interference (plus noise) is unknown. With a modest extension of the apparatus developed in this study, we can make the generalization that the signal being sought may be any vector in a given subspace, say one of dimension J, of the N space. The original problem then appears as a special case, for which $J = 1$. The physical meaning of this new signal model will depend upon the specific application, whether to radar, communications, acoustic arrays, and so on, which is being made of the theory.

The signal appears in our model as the expected value of the primary data vector z, and we now represent it as a sum:

$$E z = \sum_{j=1}^{J} b_j e_j ,$$

$$(7\text{-}1)$$

where the b_j are unknown complex amplitudes, and the vectors e_j are fixed. These vectors are assumed to be linearly independent, and their span is the J dimensional "signal subspace." Any set of vectors whose span is this subspace will suffice, and we shall see below that the LR decision rule and its performance are independent of the particular choice. It is not necessary to assume that the e_j are normalized in any special way (although that would be a harmless assumption, given the presence of the unknown amplitudes), nor need they be orthogonal.

The likelihood ratio analysis of Reference 1 is very simply generalized to include the new signal model, and we take up the derivation at the point just after the unknown covariance matrices have been estimated. The $(K+1)^{st}$ root of the likelihood ratio, which is still a function of the signal amplitudes, now has the following form:

$$\ell(b_1,...,b_J) = \frac{\|T_0\|}{\|T_1(b_1,...,b_J)\|} .$$

$$(7\text{-}2)$$

The double bars signify determinants, and

$$T_0 \equiv \frac{1}{K+1}(zz^\dagger + S)$$

$$T_1 \equiv \frac{1}{K+1}\left[\left(z - \sum_{i=1}^{J} b_i\, e_i\right)\left(z - \sum_{j=1}^{J} b_j\, e_j\right)^\dagger + S\right]$$

and

$$S \equiv \sum_{k=1}^{K} z(k)\, z(k)^\dagger \ ,$$

as before. The final LR test is then

$$\underset{b_1,\ldots,b_J}{\text{Max}}\ \ell(b_1,\ldots,b_J) = \ell = \frac{\|T_0\|}{\underset{b_1,\ldots,b_J}{\text{Min}}\ \|T_1(b_1,\ldots,b_J)\|} \geq \ell_0 \ . \tag{7-3}$$

The determinants are evaluated as in Reference 1, with the results

$$\|T_0\| = \frac{\|S\|}{(K+1)^N}\left[1 + (z^\dagger S^{-1} z)\right]$$

and

$$\|T_1(b_1,\ldots,b_J)\| = \frac{\|S\|}{(K+1)^N}\left[1 + \left(z - \sum_{i=1}^{J} b_i\, e_i\right)^\dagger S^{-1}\left(z - \sum_{j=1}^{J} b_j\, e_j\right)\right] . \tag{7-4}$$

We now define the $J \times J$ matrix

$$\Gamma_{i,j} \equiv (e_i^\dagger S^{-1} e_j) \tag{7-5}$$

and its inverse

$$\Delta \equiv \Gamma^{-1} \ . \tag{7-6}$$

Then the inner product appearing on the right side of Equation (7-4) is

58

$$\left(z - \sum_{i=1}^{J} b_i e_i\right)^{\dagger} S^{-1} \left(z - \sum_{j=1}^{J} b_j e_j\right)$$

$$= (z^{\dagger} S^{-1} z) - \sum_{i=1}^{J} b_i^* (e_i^{\dagger} S^{-1} z) - \sum_{j=1}^{J} (z^{\dagger} S^{-1} e_j) b_j + \sum_{i,j=1}^{J} b_i^* \Gamma_{i,j} b_j . \quad (7\text{-}7)$$

Next, we define the quantities

$$\beta_i \equiv \sum_{j=1}^{J} \Delta_{i,j} (e_j^{\dagger} S^{-1} z) , \quad (7\text{-}8)$$

and invert these equations to obtain

$$(e_i^{\dagger} S^{-1} z) = \sum_{j=1}^{J} \Gamma_{i,j} \beta_j .$$

When these definitions are substituted in Equation (7-7), and use is made of the evident Hermitian character of Γ, the right side of Equation (7-7) becomes

$$(z^{\dagger} S^{-1} z) + \sum_{i,j=1}^{J} \left(b_i^* \Gamma_{i,j} b_j - b_i^* \Gamma_{i,j} \beta_j - \beta_i^* \Gamma_{i,j} b_j \right)$$

$$= (z^{\dagger} S^{-1} z) + \sum_{i,j=1}^{J} (b_i - \beta_i)^* \Gamma_{i,j} (b_j - \beta_j) - \sum_{i,j=1}^{J} \beta_i^* \Gamma_{i,j} \beta_j .$$

This expression is obviously minimized by the choice

$$b_i = \beta_i ,$$

so that

$$\underset{b_1,\ldots,b_J}{\text{Min}} \| T_1(b_1,\ldots,b_J) \| = \frac{\|S\|}{(K+1)^N} \left[1 + (z^{\dagger} S^{-1} z) - \sum_{i,j=1}^{J} \beta_i^* \Gamma_{i,j} \beta_j \right] . \quad (7\text{-}9)$$

We now use this evaluation in Equation (7-3), and eliminate the β_i by means of their definition. The resulting LR test is then

$$\ell = \frac{1 + (z^\dagger S^{-1} z)}{1 + (z^\dagger S^{-1} z) - \sum_{i,j=1}^{J} (z^\dagger S^{-1} e_i) \Delta_{i,j} (e_j^\dagger S^{-1} z)} \geq \ell_0 \; . \qquad (7\text{-}10)$$

This is a direct generalization of the original LR test, as given by Equation (3-1).

Suppose we were to change to a new set of vectors for the characterization of the signal subspace, by means of a transformation such as

$$e_i' \equiv \sum_{j=1}^{J} C_{i,j} e_j \; ,$$

where C is any non-singular matrix. It is a straightforward matter to show that the form of the LR test is unchanged, inasmuch as

$$\sum_{i,j=1}^{J} (z^\dagger S^{-1} e_i) \Delta_{i,j} (e_j^\dagger S^{-1} z) = \sum_{i,j=1}^{J} (z^\dagger S^{-1} e_i') \Delta_{i,j}' (e_j'^\dagger S^{-1} z) \; , \qquad (7\text{-}11)$$

where

$$\Delta' \equiv (\Gamma')^{-1}$$

and

$$\Gamma_{i,j}' \equiv (e_i'^\dagger S^{-1} e_j') \; .$$

In this way we can, for example, convert to an orthonormal set of basis vectors for the signal subspace.

To evaluate the performance of this LR test, we follow the usual procedure and carry out a whitening transformation, assuming that the actual covariance matrix is M. We let

$$\bar{z} \equiv M^{-1/2} z$$

and

$$\bar{z}(k) = M^{-1/2} z(k) .$$

As before,

$$\bar{S} \equiv M^{-1/2} S M^{-1/2}$$

is K times the sample covariance matrix of the whitened secondaries:

$$\bar{S} = \sum_{k=1}^{K} \bar{z}(k) \bar{z}(k)^{\dagger} .$$

We also introduce the whitened vectors

$$\bar{e}_i \equiv M^{-1/2} e_i ,$$

whose span defines the signal subspace in the new system of coordinates.

Suppose the original primary vector contains the signal component

$$E z = b p , \qquad (7\text{-}12)$$

where b is a complex amplitude and p is some vector in the hypothesized signal subspace. After whitening, we will have

$$E \bar{z} = b M^{-1/2} p ,$$

which is a vector in the transformed signal subspace. This part of the discussion is the same as in Section 3, since only the signal hypothesis has changed to allow a broader class of signals.

For simplicity, we are assuming here that the actual signal conforms to the signal hypothesis. Our results can be extended directly to the mismatched case, in which the arriving signal has a component orthogonal to the hypothesized signal subspace. The implications of this generalization are discussed later on.

Inner products are unaffected by the whitening transformation, and we have

$$(z^{\dagger} S^{-1} e_i) = (\bar{z}^{\dagger} \bar{S}^{-1} \bar{e}_i)$$

61

and also

$$\Gamma_{i,j} = (\bar{e}_i^\dagger \bar{S}^{-1} \bar{e}_j) .$$

Thus,

$$\ell = \frac{1 + (\bar{z}^\dagger \bar{S}^{-1} \bar{z})}{1 + (\bar{z}^\dagger \bar{S}^{-1} \bar{z}) - \displaystyle\sum_{i,j=1}^{J} (\bar{z}^\dagger \bar{S}^{-1} \bar{e}_i) \Delta_{i,j} (\bar{e}_j^\dagger \bar{S}^{-1} \bar{z})} , \qquad (7\text{-}13)$$

and the matrix Δ is unchanged in value. Since a change of the basis vectors of the signal subspace is always possible, we may consider that the $\bar{e}_i$ have already been transformed into an orthonormal set, and no change of notation will be made. A final rotation of the data vectors can now be carried out, by means of a unitary matrix U, chosen so that the transformed basis vectors represent the first J coordinate vectors in the N space.

As in Section 3, we continue to use the old notation for these vectors, writing

$$z = U \bar{z}$$
$$z(k) = U \bar{z}(k)$$

and also

$$e_i = U \bar{e}_i .$$

The transformed primary vector now has the signal component

$$E z = b U M^{-1/2} p$$
$$= b A_p f , \qquad (7\text{-}14)$$

where f is a unit vector in the transformed signal subspace, and A_p is the same as the quantity defined in Equation (3-4). As a result of this transformation, the LR test returns to the form given in Equation (7-10), but where now

$$e_1 = [1,0,...,0]^T$$

$$e_2 = [0,1,...,0]^T$$

etc.

are the first J coordinate vectors of the N space. The matrices Γ and Δ are again given by Equations (7-5) and (7-6).

At this point we decompose the data vectors, as well as the matrix S and its inverse P, into A and B components. This decomposition differs from the one made in Section 3 in the dimensionality of the components. The A components are now J dimensional, and correspond to the signal subspace, while the B components are of dimension $N - J$. If we write

$$f = \begin{bmatrix} g \\ 0 \end{bmatrix} ,$$

where g is a unit vector, then we have

$$E z_A = b A_p g . \tag{7-15}$$

Equations (3-10) through (3-13) are still valid, and the analog of Equation (3-14) is now written

$$(z^\dagger S^{-1} z) = (z_B^\dagger S_{BB}^{-1} z_B) + y^\dagger T^{-1} y . \tag{7-16}$$

In this formula,

$$y \equiv z_A - S_{AB} S_{BB}^{-1} z_B$$

is now a J vector, and

$$T \equiv P_{AA}^{-1} = S_{AA} - S_{AB} S_{BB}^{-1} S_{BA}$$

is a $J \times J$ matrix.

63

The analysis of the matrix T follows very closely that of its scalar counterpart, given in Section 4. It begins by conditioning on the B components of the data vectors, and the matrices Q and R are then introduced by means of their old definitions. In the present case, however, the trace of R is

$$\text{Tr}[R] = K + J - N \equiv L .$$

This definition of L is consistent with our earlier usage, when J was equal to unity.

In terms of the $J \times K$ matrix

$$Z_A \equiv \left[z_A(1), \ldots, z_A(K) \right] , \tag{7-17}$$

we have

$$T = Z_A R Z_A^\dagger$$

instead of Equation (4-3). Using the identical transformation expressed by Equation (4-4), we introduce the $J \times K$ matrix

$$V_A \equiv Z_A W , \tag{7-18}$$

and then

$$T = V_A \mathscr{R} V_A^\dagger .$$

The columns of V_A are vectors which correspond to the scalars, w(m), of Equations (4-5) and (4-6). We put

$$V_A = \left[w(1), \ldots, w(K) \right] ,$$

where the w(m) are J vectors in the present case, and finally obtain

$$T = \sum_{m=1}^{L} w(m) w(m)^\dagger . \tag{7-19}$$

64

In the present application, T appears as L times the sample covariance matrix of a set of L independent Gaussian J vectors. Each of these vectors has zero mean and covariance matrix equal to the identity matrix I_J. Instead of Equation (4-8), T is now subject to a more general Wishart distribution function.

The analysis of y is quite similar, and Equations (4-9) through (4-11) require no change. The expected value of y, under the conditioning, is simply

$$E_B \, y \; = \; E \, z_A \; = \; b \, A_p \, g \; ,$$

and the covariance is now a matrix:

$$E_B \, (y - E_B \, y)(y - E_B \, y)^\dagger \; = \; [1 \; + \; (z_B^\dagger S_{BB}^{-1} z_B)] \, I_J \; .$$

The proof of the independence of y and T, given in Section 4, applies directly to the present more general situation.

The analog of v is now a J vector:

$$v \; = \; [1 \; + \; (z_B^\dagger S_{BB}^{-1} z_B)]^{-1/2} \, y \; , \tag{7-20}$$

whose conditional covariance matrix is just I_J, and whose conditional mean is

$$E_B \, v \; = \; b \, A_p \, \rho^{1/2} g \; . \tag{7-21}$$

We have introduced the loss factor

$$\rho \; \equiv \; \frac{1}{1 \; + \; (z_B^\dagger S_{BB}^{-1} z_B)} \; , \tag{7-22}$$

which now depends on the $N - J$ dimensional secondary data vectors. In the present case, we have the correspondences

$$(z_B^\dagger S_{BB}^{-1} z_B) \; = \; \Sigma(N - J, K)$$

and

$$\rho \; = \; \rho(N - J, K) \; . \tag{7-23}$$

65

When these results are combined, we see that Equation (7-16) can be written

$$1 + (z^\dagger S^{-1} z) = [1 + (z_B^\dagger S_{BB}^{-1} z_B)] [1 + (v^\dagger T^{-1} v)] , \tag{7-24}$$

in which the second factor on the right is conditioned on the B components only through the loss factor ρ. Since v and T are independent, Equation (7-24) is equivalent to the factorization expressed by

$$1 + \Sigma(N,K) = [1 + \Sigma(N-J,K)] [1 + \Sigma(J,K+J-N)] . \tag{7-25}$$

If we replace J by $N - M$, Equation (7-25) is the same as Equation (4-29), which was inferred by induction in Section 4.

These results apply directly to the analysis of the LR test, expressed by Equation (7-10). Since the e_j are now the basis vectors of the signal subspace, we see that the matrix Γ, whose elements are defined by Equation (7-5), is identical to P_{AA}. It is also clear that the quantities

$$(e_i^\dagger S^{-1} z)$$

are simply the components of the J vector

$$P_{AA} z_A + P_{AB} z_B = P_{AA} y .$$

Thus,

$$\sum_{i,j=1}^{J} (z^\dagger S^{-1} e_i) \Delta_{i,j} (e_j^\dagger S^{-1} z) = y^\dagger P_{AA} y = y^\dagger T^{-1} y ,$$

and hence, according to Equation (7-16), the LR test becomes

$$l = \frac{1 + (z^\dagger S^{-1} z)}{1 + (z_B^\dagger S_{BB}^{-1} z_B)} \geq l_0 , \tag{7-26}$$

a direct generalizaton of Equation (3-20).

Application of the factorization identity expressed by Equation (7-24) yields the form

$$(v^\dagger T^{-1} v) \geq \ell_0 - 1$$

for the LR test. Since the left side of this equation is identical in structure to the quantity $\Sigma(J, K + J-N)$, already studied in Section 5, we can apply the results of that section directly to obtain the performance of the generalized LR test in terms of its PD and PFA.

We must retain the conditioning on the B components at first, so that v will be Gaussian, with expected value given by Equation (7-21). The conditional probability of detection will then be given by Equation (5-24), with the following parameter assignments:

$$\mathcal{N} = J$$
$$\mathcal{L} = K+1-N = L+1-J$$
$$c = |b|^2 A_p^2 \rho$$
$$\xi = \ell_0 - 1 .$$

(7-27)

We define the signal parameter

$$a \equiv |b|^2 A_p^2$$

(7-28)

which is consistent with our former definition, Equation (4-19), since the present situation corresponds to $\Theta = 0$. Substituting, we obtain for the conditional probability of detection

$$P_D(B) = 1 - \frac{1}{\ell_0^L} \sum_{m=J}^{L} \binom{L}{m}(\ell_0-1)^m G_{m+1-J}\left(\frac{a\rho}{\ell_0}\right)$$

(7-29)

According to Equations (7-23) and (5-34), the PDF of the present loss factor ρ will be simply

$$f(\rho) = f(\rho;N-J,K,0) = f_\beta(\rho;K+1+J-N,N-J) = f_\beta(\rho;L+1,N-J) ,$$

since there are no signal components in the secondaries in this problem. If we wished to extend the present analysis to the mismatched case, we would postulate the presence of a signal in the primary vector which would have a component outside the signal subspace. The effect would be to add a signal to z_B in the present context, and then the PDF of the loss factor ρ would be given by the more general form of Equation (5-34), with an appropriate value for the signal parameter c.

The computation of the detection probability of the generalized LR test is completed by the removal of the conditioning, which again is confined to the loss factor. The result is

$$P_D = 1 - \frac{1}{\ell_0^L} \sum_{m=J}^{L} \binom{L}{m} (\ell_0 - 1)^m H_{m+1-J}\left(\frac{a}{\ell_0}\right),$$ (7-30)

where

$$H_m(y) \equiv \int_0^1 G_m(\rho y) f_\beta(\rho; L+1, N-J) \, d\rho$$ (7-31)

and, of course,

$$L = K + J - N.$$

The corresponding PFA is obtained by putting $a = 0$. Since $G_m(0) = 1$, Equation (7-31) gives $H_m(y) = 1$, and the PFA becomes

$$PFA = \frac{1}{\ell_0^L} \sum_{m=0}^{J-1} \binom{L}{m} (\ell_0 - 1)^m.$$ (7-32)

When $J = 1$, this PFA is a single term, already given by Equation (4-26). Equation (7-30) reduces immediately to Equation (4-23) when $J = 1$, and Equation (7-31) is the same as Equation (4-24) when the appropriate PDF is used for the loss factor in that equation. This loss factor will correspond to the absence of signals in the secondaries, i.e., $\theta = 0$. Expressions suitable for numerical computation of the detection probability, both for the problem analyzed in this section and the general problem addressed in Sections 3, 4 and 5, are given in the Appendix.

68

8. SUBSPACE SIGNALS; NUMERICAL RESULTS AND DISCUSSION

In the problem of detecting a signal which is hypothesized to lie in a subspace of the observation space, the basic parameters are the assigned PFA, the signal strength SNR_{pp}, and the dimensionality of the subspace J. This signal strength parameter was defined in Equation (6-1) as the "available SNR," and it is the only SNR parameter required in the present instance, since the actual signals have been assumed to lie within the specified subspace. In other words, there is no analog of the Θ parameter. It should be kept in mind that a subspace of the N dimensional observation space does not correspond to a subset of directions of arrival of signals in real space, except in very special cases.

An explicit formula for the detection probability is derived in the Appendix. It is simpler than the analogous formula for the mismatch problem, because the loss factor PDF in the present case is a simple Beta distribution. Detection performance is illustrated in Figure 8-1, which shows PD as a function of SNR for the case N = 4 and K = 20. The curves are parametrized by all possible values of J.

When J = 4 the "subspace" is the entire observation space, and the decision rule becomes a test between the hypothesis that the primary sample vector has expected value zero and the hypothesis that its expected value is any non-zero vector. No signal structure is postulated in this case but the total interference is still modeled as having zero mean and unknown covariance. We note that at the 0.9 level of detection probability, the penalty for using this more general hypothesis (i.e., J = N) is slightly more than 1 dB, in the case illustrated in Figure 8-1.

Such a test could be useful in a rapid search for signal energy in a situation where signals are expected to be scarce. It could then be followed up with more conventional tests in those antenna pointing directions for which signals were found. The general subspace test, of course, gives no indication of the direction of a detected signal within the designated subspace.

A similar family of surves is presented in Figure 8-2 for the case of a 20-dimensional subspace, using 100 secondary vectors. Selected values of J have been chosen for the curves. The penalty (at a PD of 0.9) of testing for a totally unstructured signal is now about 3 dB.

To understand the operation of the LR test in this application, we recast the decision rule, using the inner product notation introduced in Section 6. Equation (7-10), which expresses this rule, can then be written
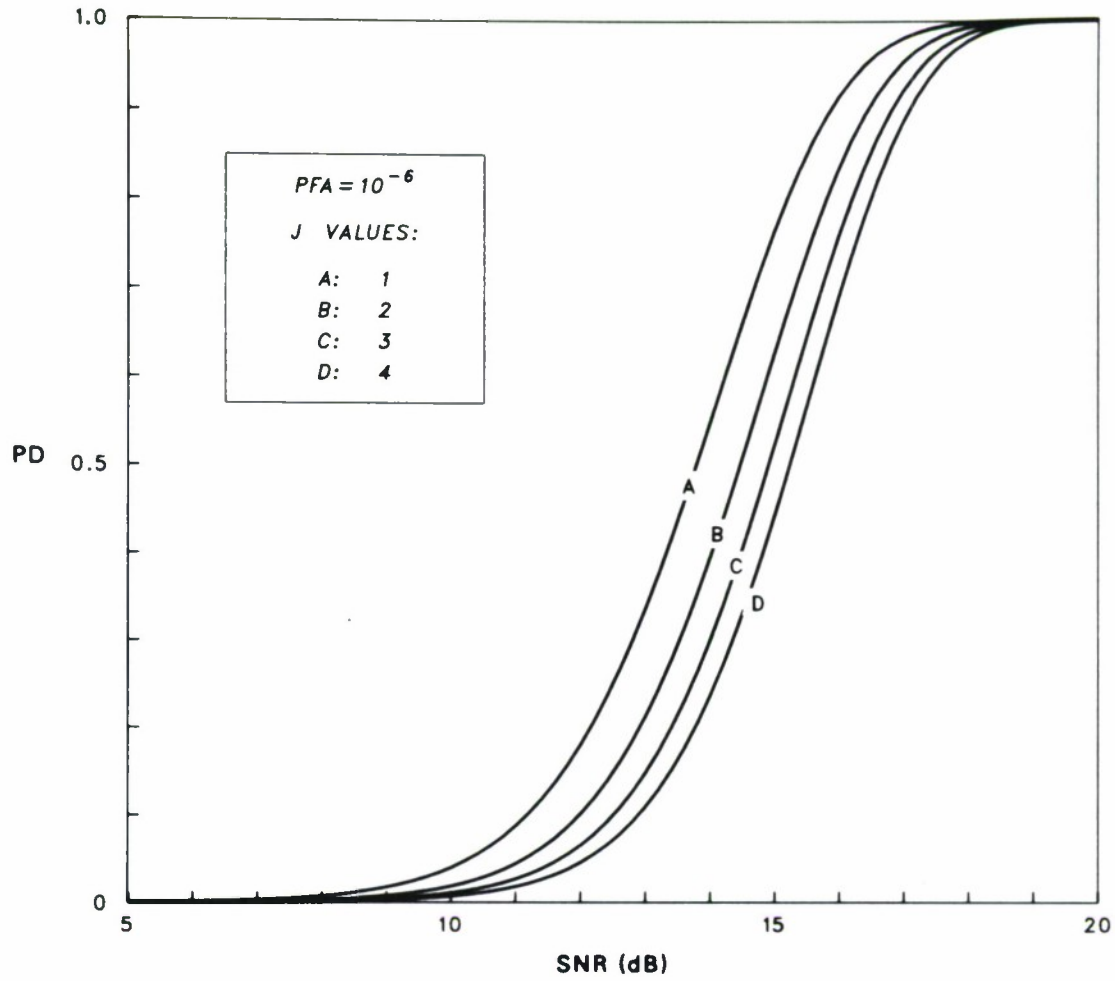
69

Figure 8-1. Subspace signal probability of detection; N = 4 and K = 20.

$$\frac{K + [z, z]}{K + [z, z] - \sum_{i, j=1}^{J} [z, e_i] \frac{1}{K} \Delta_{i,j} [e_j, z]} \geq \ell_0 \,, \tag{8-1}$$

where the e_i represent an arbitrary linearly-independent set of vectors whose span is the desired subspace. From their definitions we also have

$$\frac{1}{K} \Delta = (K \Gamma)^{-1}$$
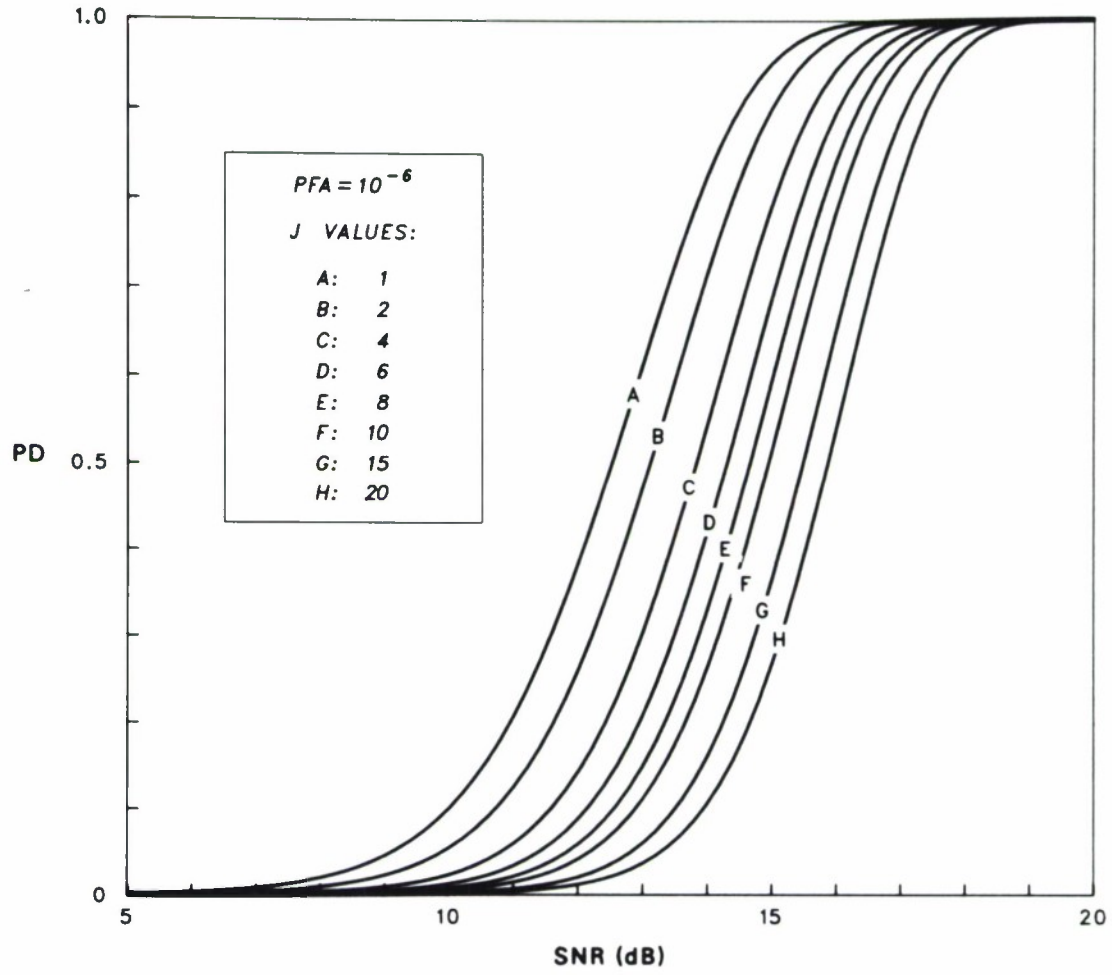
70

Figure 8-2. Subspace signal probability of detection; N = 20 and K = 100.

and

$$K \, \Gamma_{i,j} = [e_i, e_j] \ .$$

It is easy to show that the double sum

$$\frac{1}{K} \sum_{i,j=1}^{J} e_i \, \Delta_{i,j} \, e_j^\dagger$$

71

is a projection operator into the subspace, and that the corresponding term in Equation (8-1) is the norm of that component of z which lies in this subspace. Approaching this by a different route, we introduce a new set of vectors to define the subspace, by means of a linear transformation on the e_i, chosen so that the new vectors form an orthonormal set. As pointed out in Section 7 [see Equation (7-11)], the form of the LR test is unchanged by such a transformation. The matrix Δ, however, becomes K times the identity matrix in this special set of basis vectors, and the test itself becomes

$$\frac{K + [z,z]}{K + [z,z] - \sum_{i=1}^{J} |[e_i,z]|^2} \geq \ell_0 \ . \tag{8-2}$$

Finally, we define the vector

$$z_{\parallel} \equiv \sum_{i=1}^{J} [e_i,z] e_i \ , \tag{8-3}$$

which is the component of z in the subspace, and its orthogonal complement

$$z_{\perp} \equiv z - z_{\parallel} \ . \tag{8-4}$$

These components are simple generalizations of the quantities given these names in Section 6. With the help of this notation, we see that the LR test can now be written in the identical form which was found in Section 6, namely:

$$\frac{[z_{\parallel},z_{\parallel}]}{1 + \frac{1}{K}[z_{\perp},z_{\perp}]} \geq K(\ell_0 - 1) \ . \tag{8-5}$$

If K is allowed to increase without limit, the LR test passes over into the test

$$\sum_{i=1}^{J} |[e_i,z]|^2 \geq \text{constant} \ ,$$

or

$$\sum_{i=1}^{J} |(e_i^\dagger M^{-1} z)|^2 \geq \text{constant} \tag{8-6}$$

in the original notation. The e_i here are orthonormal basis vectors of the subspace.

This asymptotic test has a simple significance. Each term in the sum is the decision statistic for the detection of a signal in the direction of a basis vector, and the final test is a noncoherent summation of these quantities. This makes good sense as a way of testing for a signal whose structure only places it in a subspace, and the LR test for finite K is a CFAR version of this principle which operates with unknown interference.

One expects the noncoherent combination to bring with it a performance penalty, just as does conventional noncoherent integration with respect to coherent integration. In the present case, where a performance penalty is already being paid for interference estimation and CFAR operation, it appears that the additional penalty for a broadening of the signal hypothesis is not great. A somewhat similar situation arises with fluctuating targets of a more familiar kind, namely that noncoherent integration is more efficient than with nonfluctuating targets, since the fluctuation loss is being overcome at the same time. In our case, the loss factor represents a kind of fluctutation phenomenon, and perhaps this explains the modest loss associated with subspace detection.

APPENDIX: EVALUATION OF THE PROBABILITY OF DETECTION

The probability of detection for the generalized problem of Section 7 was given in the form

$$P_D = 1 - \frac{1}{\ell_0^L} \sum_{m=J}^{L} \binom{L}{m} (\ell_0 - 1)^m H_{m+1-J}\left(\frac{a}{\ell_0}\right) , \qquad (A\text{-}1)$$

where

$$L = K + J - N ,$$

a is the signal parameter, and

$$H_m(y) \equiv \int_0^1 G_m(\rho y) f(\rho) \, d\rho . \qquad (A\text{-}2)$$

In this integral, $f(\rho)$ is the loss factor PDF and

$$G_m(y) \equiv e^{-y} \sum_{k=0}^{m-1} \frac{y^k}{k!} . \qquad (A\text{-}3)$$

The corresponding PFA is

$$PFA = \frac{1}{\ell_0^L} \sum_{m=0}^{J-1} \binom{L}{m} (\ell_0 - 1)^m . \qquad (A\text{-}4)$$

In the problem of detecting a signal hypothesized to lie in a subspace, we assumed that the actual signal conforms to this hypothesis. In that case, the signal parameter is

$$a = |b|^2 A_p^2 = |b|^2 (p^\dagger M^{-1} p) . \qquad (A\text{-}5)$$

The actual interference covariance matrix is taken to be M, the unit vector p is in the direction of the actual signal, and b is a complex signal amplitude.

For the problem of the mismatched signal (discussed in Sections 3, 4, and 5), the signal parameter is

$$a = |b|^2 A_p^2 \cos^2 \Theta = |b|^2 \frac{|(q^\dagger M^{-1} p)|^2}{(q^\dagger M^{-1} q)} , \qquad (A-6)$$

where q is the steering direction and p is again a unit vector in the direction of the actually arriving signal. In this case we also have $J=1$, hence the detection and false alarm probabilities are

$$P_D = 1 - \frac{1}{\ell_0^L} \sum_{m=1}^{L} \binom{L}{m} (\ell_0 - 1)^m H_m\left(\frac{a}{\ell_0}\right) \qquad (A-7)$$

and

$$PFA = \frac{1}{\ell_0^L} , \qquad (A-8)$$

where now

$$L = K + 1 - N .$$

It is a simple matter to solve Equation (A-8) for the threshold, given the false alarm probability. For the subspace problem, however, Equation (A-4) must be inverted, and this can be accomplished by means of the Newton-Raphson iteration. The threshold obtained by first solving Equation (A-8) can be used as a starting value. The iteration itself requires the derivative of PFA with respect to the threshold parameter, and this is given by the formula

$$\frac{d\,PFA}{d\ell_0} = - \frac{L}{\ell_0^{L+1}} \binom{L-1}{J-1} (\ell_0 - 1)^{(J-1)} ,$$

which can be verified in a straightforward manner, making use of the binomial identity given in Equation (5-26).

To proceed with the evaluation, we substitute Equation (A-3) in Equation (A-2), and make a change of the variable of integration, replacing ρ by $1 - \rho$. The result is

$$H_m(y) = \sum_{k=0}^{m-1} \frac{y^k}{k!} \int_0^1 e^{-\rho y} f(\rho) \rho^k \, d\rho$$

$$= e^{-y} \sum_{k=0}^{m-1} \frac{y^k}{k!} J_k(y) , \qquad (A\text{-}9)$$

where the new functions are

$$J_k(y) = \int_0^1 e^{\rho y} f(1-\rho)(1-\rho)^k \, d\rho . \qquad (A\text{-}10)$$

The change of variable is introduced to obtain a series of exclusively positive terms, which is important to ensure the stability of the numerical computation. The evaluation of detection probability can be carried out in a variety of ways, leading to very different expressions as multiple series. The method employed here, which is a direct generalization of that used in Reference 1, provides a practical basis for computation.

If we define

$$H_0(y) \equiv 0 ,$$

then Equation (A-9) is equivalent to the recursion relation

$$H_m(y) = H_{m-1}(y) + e^{-y} \frac{y^{m-1}}{(m-1)!} J_{m-1}(y) , \qquad (A\text{-}11)$$

which is easily programmed for computation.

We proceed with the evaluation of the $J_m(y)$, using the PDF of the general loss function derived in Section 5;

$$f(\rho) = f(\rho; \mathcal{N}, K, c) = e^{-c\rho} \sum_{s=0}^{\mathscr{L}} \binom{\mathscr{L}}{s} \frac{(\mathcal{N} + \mathscr{L} - 1)!}{(\mathcal{N} + \mathscr{L} - 1 + s)!} c^s f_\beta(\rho; \mathscr{L}, \mathcal{N} + s) , \qquad (A\text{-}12)$$

where

$$\mathcal{L} \equiv K+1-\mathcal{N}$$

and c is a signal parameter. This will provide answers to both problems by specialization of the parameters. For the mismatch problem, we have

$$\mathcal{N} = N-1$$
$$\mathcal{L} = K+2-N = L+1$$
$$c = |b|^2 A_p^2 \sin^2 \Theta \,, \tag{A-13}$$

while for the generalized problem of Section 7, we must put

$$\mathcal{N} = N-J$$
$$\mathcal{L} = K+1+J-N = L+1$$
$$c = 0 \,. \tag{A-14}$$

We substitute Equation (A-12) in Equation (A-10):

$$J_m(y) = \sum_{s=0}^{\mathcal{L}} \binom{\mathcal{L}}{s} \frac{(\mathcal{N}+\mathcal{L}-1)!}{(\mathcal{N}+\mathcal{L}-1+s)!} c^s \int_0^1 e^{\rho y} e^{-c(1-\rho)} f_\beta(1-\rho;\mathcal{L},\mathcal{N}+s)(1-\rho)^m \, d\rho$$

$$= e^{-c} \sum_{s=0}^{\mathcal{L}} \binom{\mathcal{L}}{s} \frac{(\mathcal{N}+\mathcal{L}-1)!}{(\mathcal{N}+\mathcal{L}-1+s)!} c^s \int_0^1 e^{\rho(y+c)} f_\beta(\rho;\mathcal{N}+s,\mathcal{L})(1-\rho)^m \, d\rho \,.$$

In the second line we have made use of an obvious symmetry of the Beta distribution. Next we substitute the definition of this Beta function, and absorb the power of $1-\rho$ into a new Beta function, with the result

$$J_m(y) = e^{-c} \sum_{s=0}^{\mathcal{L}} \binom{\mathcal{L}}{s} \frac{(\mathcal{N}+\mathcal{L}-1)!(\mathcal{L}-1+m)!}{(\mathcal{N}+\mathcal{L}-1+m+s)!(\mathcal{L}-1)!} c^s \int_0^1 e^{\rho(y+c)} f_\beta(\rho;\mathcal{N}+s,\mathcal{L}+m) \, d\rho \,. \tag{A-15}$$

By expanding the exponential, integrating term by term and recognizing the result as a hypergeometric function, one easily obtains the formula

$$\int_0^1 e^{z\rho} f_\beta(\rho;n,m)\,d\rho = {}_1F_1[n;n+m;z] \ .$$

This is a well known integral representation for the confluent hypergeometric function (see Reference 7), which appears here as a moment generating function for the Beta distribution. Applied to Equation (A-15), we obtain the desired result

$$J_m(y) = e^{-c} \sum_{s=0}^{\mathscr{L}} \binom{\mathscr{L}}{s} \frac{(\mathscr{N}+\mathscr{L}-1)!\,(\mathscr{L}-1+m)!}{(\mathscr{N}+\mathscr{L}-1+m+s)!\,(\mathscr{L}-1)!}\, c^s\ {}_1F_1[\mathscr{N}+s;\mathscr{N}+\mathscr{L}+m+s;y+c] \ . \qquad \text{(A-16)}$$

When the signal parameter vanishes, this formula simplifies to one term:

$$J_m(y) = \frac{(\mathscr{N}+\mathscr{L}-1)!\,(\mathscr{L}-1+m)!}{(\mathscr{N}+\mathscr{L}-1+m)!\,(\mathscr{L}-1)!}\ {}_1F_1[\mathscr{N};\mathscr{N}+\mathscr{L}+m;y] \ . \qquad \text{(A-17)}$$

The detection probability formula derived in Reference 1 agrees with Equation (A-17), when the parameter assignments of Equation (A-13) are used, since the analysis of that reference did not include signal mismatch.

When the signal amplitude becomes very large, both y and c tend to infinity, and the asymptotic approximation for the confluent hypergeometric function (Reference 7) can be applied to Equation (A-16). The result is that

$$e^{-y}\,\frac{y^m}{m!}\,J_m(y) \to \frac{(\mathscr{L}+m-1)!}{m!\,(\mathscr{L}-1)!}\ \frac{y^m\,c^{\mathscr{L}}}{(y+c)^{\mathscr{L}+m}} \ , \qquad \text{(A-18)}$$

in this limit. This formula has been used to obtain the asymptotic detection curves that appear in Section 6. Equation (A-18) can also be obtained by first evaluating the asymptotic form of the loss factor PDF, going back to Equation (5-27). Retaining only the term $m=\mathscr{L}$ in that series, and substituting in Equation (A-10), the same expression for $J_m(y)$ can be derived.

Equations (A-1), (A-9), and (A-16) or (A-17) are the basis for the evaluation of the detection probability curves of Sections 6 and 8. In the confluent hypergeometric

functions, the second argument exceeds the first, which results in monotonically decreasing coefficients in these series. It is then easy to obtain truncation bounds, in close analogy to the work of Shnidman[8] on the Marcum Q-function. Shnidman's technique for coping with underflow has also been found necessary here, to get answers for a wide range of the parameters.

REFERENCES

1. E.J. Kelly, "Adaptive Detection in Non–Stationary Interference, Part I and Part II," Technical Report 724, Lincoln Laboratory, MIT (25 June 1985), DTIC AD-A158810.

2. D.M. Boroson, "Sample Size Considerations for Adaptive Arrays," IEEE Trans. Aerosp. Electron. Syst. **AES-16**, 446 (1980).

3. I.S. Reed, J.D. Mallett, and L.E. Brennan, "Rapid Convergence Rate in Adaptive Arrays," IEEE Trans. Aerosp. Electron. Syst. **AES-10**, 853 (1974).

4. M.D. Springer, *The Algebra of Random Variables* (Wiley, New York, 1979).

5. Private communication.

6. E.J. Kelly, "Finite-Sum Expressions for Signal Detection Probabilities," Technical Report 566, Lincoln Laboratory, MIT (20 May 1981), DTIC AD-A102143.

7. E.T. Copson, *An Introduction to the Theory of Functions of a Complex Variable* (Oxford University Press, London, 1935).

8. D.A. Shnidman, "Efficient Evaluation of Probabilities of Detection and the Generalized Q-Function," IEEE Trans. Inf. Theory **IT-22**, 746 (1976).

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY 2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
4. PERFORMING ORGANIZATION REPORT NUMBER(S) Technical Report 761	5. MONITORING ORGANIZATION REPORT NUMBER(S) ESD-TR-86-090

6a. NAME OF PERFORMING ORGANIZATION Lincoln Laboratory, MIT	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Electronic Systems Division
6c. ADDRESS (City, State, and Zip Code) P.O. Box 73 Lexington, MA 02173-0073		7b. ADDRESS (City, State, and Zip Code) Hanscom AFB, MA 01731
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Air Force Systems Command, USAF	8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F19628-85-C-0002

8c. ADDRESS (City, State, and Zip Code) Andrews AFB Washington, DC 20334	10. SOURCE OF FUNDING NUMBERS			
	PROGRAM ELEMENT NO. 63250F	PROJECT NO. 227	TASK NO.	WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)

Adaptive Detection in Non-stationary Interference, Part III

12. PERSONAL AUTHOR(S)
Edward J. Kelly

13a. TYPE OF REPORT Technical Report	13b. TIME COVERED FROM _____ TO _____	14. DATE OF REPORT (Year, Month, Day) 1987, August, 24	15. PAGE COUNT 92

16. SUPPLEMENTARY NOTATION

None

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	adaptive detection maximum-likelihood estimation statistical hypothesis testing

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The analysis of Parts I and II of the report with this title has been extended in two directions. In the first case, the performance of an adaptive system with respect to signals arriving from directions other than the steering direction is evaluated. It is shown that these signals are rejected more strongly than would be suggested by the sidelobe levels of the adaptive patterns themselves. In the other case, the detection problem is generalized to include the detection of signals known only to lie in a subspace of the space of steering vectors. Again, performance is derived and the penalty associated with the greater uncertainty of the signal model is shown to be small. The analysis of Part I is essentially repeated here, both to keep this report self-contained and to present an alternative version of the basic derivations.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Maj. Thomas J. Alpert, USAF	22b. TELEPHONE (Include Area Code) (617) 863-5500, x-2330 22c. OFFICE SYMBOL ESD/TML

DD FORM 1473, 84 MAR
83 APR edition may be used until exhausted.
All other editions are obsolete.